

Discerning Non-Stationary Market Microstructure Noise and Time-Varying Liquidity in High Frequency Data*

Richard Y. Chen[†], Per A. Mykland[‡]

October 2015

Abstract

In this paper, we investigate the implication of non-stationary market microstructure noise to integrated volatility estimation, provide statistical tools to test stationarity and non-stationarity in market microstructure noise, and discuss how to measure liquidity risk using high frequency financial data. In particular, we discuss the impact of non-stationary microstructure noise on TSRV (Two-Scale Realized Variance) estimator, and design three test statistics by exploiting the edge effects and asymptotic approximation. The asymptotic distributions of these test statistics are provided under both stationary and non-stationary noise assumptions respectively, and we empirically measure aggregate liquidity risks by these test statistics from 2006 to 2013. As byproducts, functional dependence and endogenous market microstructure noise are briefly discussed. Our empirical study indicates the prevalence of non-stationary market microstructure noise in the New York Stock Exchange.

Keywords: Market Microstructure Noise, Integrated and Realized Volatilities, Non-stationarity, Edge Effects, Testing, Stable Central Limit Theorems, Liquidity

JEL classification: C12, C13, C14, C58

*The authors are grateful for support under National Science Foundation Grant DMS 14-07812. All comments are gratefully welcomed.

[†]Richard Y. Chen is a current Ph.D. student in Statistics at the University of Chicago. Email: yrchen@uchicago.edu.

[‡]Per A. Mykland is the Robert M. Hutchins Distinguished Service Professor, Department of Statistics, the University of Chicago. Email: mykland@pascal.uchicago.edu.

1 Introduction

The introduction of high-tech trading mechanisms into markets, for example, electronic communication networks (ECNs) and other electronic trading platforms, provides an opportunity for speculators and market makers to take advantage of speed in trading and market making, and this technological innovation also brings new regulatory challenges. The subsequent high-frequency trading results in a huge amount of highly frequently observed financial data, which, in particular, open two potential gates for research in theoretical and empirical asset pricing: one is estimation methodology using high-frequency data, since practitioners and researchers can get access to the big data and estimate variables of interest with greater accuracy; the other is a “frog eyes’ view” on market microstructure, since low-latency data can offer a chance to investigate the market’s trading behaviors with a higher resolution than ever before.

Correspondingly, this paper’s contributions to the literature are twofold: *i)* one is testing non-stationary market microstructure noise, we study the estimation problem when using high-frequency data with non-stationary noises, and then test non-stationarity in market microstructure noise via edge effect; *ii)* the other one is on empirical market microstructure, where we estimate the noise as measures of time-varying bid-ask spreads, risk aversions of market participants etc, and detect short-term liquidity changes.

1.1 Literature review

The high-frequency finance practice motivates two clearly distinct and closely related researches:

One is more accurate estimation in financial econometrics, to name a few but not all, the estimation of integrated volatilities, quadratic covariances, the activities of jumps, the leverage effects, the volatility of volatility, the lead-lag effect. This stream of research started from [Jacod \[1994\]](#), [Jacod and Protter \[1998\]](#) from the perspective of stochastic calculus, and [Foster and Nelson \[1996\]](#), [Engle \[2000\]](#), [Zhang \[2001\]](#), [Andersen et al. \[2001\]](#), [Barndorff-Nielsen and Shephard \[2002\]](#) in the context of econometrics. Now the high-frequency financial econometrics has already developed into a considerably influential research field with numerous prominent scholars and there are already monographs on this area: [Jacod and Shiryaev \[2003\]](#), [Jacod and Protter \[2012\]](#) developed probabilistic tools for high-frequency financial data analysis, [Aït-Sahalia and Jacod \[2014\]](#) provided an excellent overview in econometric literature, [Hautsch \[2012\]](#) is a good account from a financial standpoint. There are also academic chapters concisely reviewing high-frequency financial econometrics: [Russell and Engle \[2010\]](#), [Mykland and Zhang \[2012\]](#), [Jacod \[2012\]](#).

The other one is the study of market microstructure. The low-latency data allow financial practitioners and researchers to look at the financial markets at a higher resolution level, for example, one can know the bid/ask dynamics within each second, one can also know the order flow through the limit order book. The market microstructure theory studies how the latent demand and latent supply of market participants are ultimately translated into prices by studying the specific market structure in de-

tail. The cornerstone papers include [Glosten and Milgrom \[1985\]](#), [Kyle \[1985\]](#), both of them are using (pesudo)¹ game-theoretical argument in information economics. More comprehensive books include [O’Hara \[1998\]](#), [Hasbrouck \[2007\]](#). However, when looking closely at the transaction or quotation prices, one can find that the price is no longer a semimartingale, not even random walk. For this reason, according to market microstructure theory [[O’Hara, 2003](#)], the semimartingale model in classical asset pricing theory [[Harrison and Pliska, 1981](#), [Delbaen and Schachermayer, 1994](#)] is not a photographic depiction of the real prices of financial assets, yet it is still a fairly good approximation to asset prices when the trading frequency is not sufficiently high, and that is the reason the literature suggests using 5-minute subsampling.

Some estimation methods for integrated volatility using noisy high-frequency financial data have already been well established: *i)* [Zhang et al. \[2005\]](#) found the first consistent estimator (two-time scale realized volatility) using subsampling and averaging in the presence of i.i.d. market microstructure noise and [Zhang \[2006\]](#) gave a multi-scale version with the optimal rate of convergence $n^{\frac{1}{4}}$, [Li and Mykland \[2007\]](#) discussed the robustness of TSRV to noise assumptions in general, [Kalnina and Linton \[2008\]](#) generalized the TSRV to the model with endogeneous and dirual noise and put forward a modified version of TSRV which we shall use in this paper. Later [Aït-Sahalia et al. \[2011\]](#) generalized the model to allowing correlated noises under stationary and strong-mixing conditions; *ii)* [Barndorff-Nielsen et al. \[2008\]](#) provided a kernel-based estimator under the model in which the noise process is temporarily dependent and stationary and possibly linearly correlated with the latent Itô process, their inference is also robust to endogeneous spacing; *iii)* [Jacod et al. \[2009\]](#) designed a generalized version of the pre-averaging approach [[Podolskij and Vetter, 2009](#)], under a Markovian noise model which allows arbitrary fashion of noise but without correlation between noise and the latent process; *iv)* Motivated by the likelihood method from [Aït-Sahalia et al. \[2005\]](#), [Xiu \[2010\]](#) established quasi-maximum likelihood method (QMLE) in the estimation of integrated volatility; *v)* [Bibinger et al. \[2014\]](#) developed the local generalized method of moments to estimate quadratic covariation using noisy high-frequency data.

Many estimators of integrated volatilities using high-frequency noisy data were developed under the assumption that the microstructure noise is stationary. As argued by [Aït-Sahalia et al. \[2006, 2011\]](#), without stationarity, time-dependent noise components might not even be identifiable. However, literature in empirical finance, such as [Admati and Pfleiderer \[1988\]](#), [Hasbrouck \[1993\]](#), [Andersen and Bollerslev \[1997\]](#), [Gouriérous et al. \[1999\]](#), has already shown in 1990s that markets exhibit a systematic intra-day pattern of regular variations. Therefore, allowing heteroskedasticity and non-stationary in market microstructure noise in integrated volatility estimation is of particular importance in application. Particularly, [Kalnina and Linton \[2008\]](#) used a parametric model to describe the diurnal pattern in microstructure noise.

Besides, [Aït-Sahalia and Yu \[2009\]](#) used the estimates of noise variance in high-frequency data to measure the market liquidity from June 1996- December 2005. There

¹To say it “pseudo” because in model considered in [Kyle \[1985\]](#), the market maker does not aim to maximize their utility, instead his or her objective is only to guarantee market clearing.

is other related research in the literature, Awartani et al. [2009] studied the changes in microstructure noise due to sampling frequency, Bandi et al. [2013] derived the optimal sampling frequency in terms of finite-sample forecast mean squared error in linear forecast model with non-stationary market microstructure noise.

Accordingly, this paper contributes to the literature in two ways. On one hand, we study the impact of non-stationary noise on integrated volatility estimation, in particular, we utilize the edge-effect correction in Kalnina and Linton [2008] for the TSRV of Zhang et al. [2005], Li and Mykland [2007], and get the similar asymptotic result in a more general model of the noise process motivated by Li and Mykland [2007], Jacod et al. [2009] (similar models are in Jacod et al. [2010] and Chapter 16 of Jacod and Protter [2012]). Our main result is that we could test the existence of non-stationary noise by exploiting the edge effects in TSRV due to the non-stationary noise. On the other hand, based on our test statistic, we could compare the second moments of market microstructure noises across different time periods, and evaluate the short-term liquidity shifts in the financial markets by our test statistics since the market microstructure noise can capture some information about market quality and liquidity [Hasbrouck, 1993, Stoll, 2000, O'Hara, 2003, Aït-Sahalia and Yu, 2009].

1.2 Structure of the paper

This article is organized in the following way:

In Section 2, we describe a model with general noise structure; in Section 3, we will discuss the edge effect in the original TSRV due to non-stationary market microstructure noises, and prove the satisfactory asymptotic property of “sample-weighted” TSRV based on previous calculation;

In Section 4 and 5, we design some statistical tests based on in-fill asymptotic approximation and edge effects due to the non-stationarity to test whether the market microstructure noise is indeed stationary or not. We also provide their stable central limit theorems under the null hypothesis;

Section 6 first introduce an aggregate measure of liquidity risks, then investigate the behaviors of our test statistics in Section 5 when the market microstructure noise is time-varying (stable central limit theorems are provided), based on which we can study the power of the test statistics as well as estimation problem for the aggregate liquidity risks.

In Section 7, we further discuss how to make inference about the relation between the latent process $\{\sigma_t^2\}_{t \geq 0}$ and the conditional second moment of noise $\{\epsilon_t\}_{t \geq 0}$ non-parametrically in 7.1; and we will discuss an extension of the general model in 7.2, which allows endogenous noise inspired by market microstructure theory;

In Section 8 and 9, we conduct some simulation studies to corroborate the theoretical study, and conduct empirical analysis using DJIA30 data in 9.1 and 9.2, using our tests, we show the prevalence of non-stationary market microstructure noises in the U.S. stock market, and find a striking serial pattern of liquidity;

In Section 10, we draw our conclusions. Proofs of the lemmas and theorems are provided in the Appendix (Section 11).

2 The model and assumptions

In this article, we consider a general model with arbitrary fashion of noise (including additive noise, round-off error, thereof combined, and others). The setup is the same as the model in Jacod et al. [2009].

2.1 Model setup

Firstly, we have a filtered probability space $\left(\Omega^{(0)}, \mathcal{F}^{(0)}, \{\mathcal{F}_t^{(0)}\}_{t \geq 0}, \mathbb{P}^{(0)}\right)$ on which the latent Itô semimartingale $\{X_t\}_{t \geq 0}$ is defined. The Itô semimartingale $\{X_t\}_{t \geq 0}$ can be described by:

$$X_t = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s \quad (1)$$

where $\{b_t\}_{t \geq 0}$ and $\{\sigma_t\}_{t \geq 0}$ are càdlàg processes and adapted to $\{\mathcal{F}_t^{(0)}\}_{t \geq 0}$, σ_t^2 is the volatility in financial terminology (for example, it can be described by the Heston model [Heston, 1993]), $\{W_t\}_{t \geq 0}$ is a standard 1-dimensional Wiener process.

Secondly, we have another filtered probability space on which the observable process $\{Y_t\}_{t \geq 0}$ is defined: $\left(\Omega^{(1)}, \mathcal{F}^{(1)}, \{\mathcal{F}_t^{(1)}\}_{t \geq 0}, \mathbb{P}^{(1)}\right)$. Then we can define the market microstructure noise process, $\{e_t\}_{t \geq 0}$ ², as the difference between the latent and observable processes:

$$e_t = Y_t - X_t \quad (2)$$

Besides we define:

$$Z_t \equiv E_{\mathbb{P}^{(1)}}(Y_t | \omega^{(0)}) = X_t + E_{\mathbb{P}^{(1)}}(e_t | \omega^{(0)}) \quad (3)$$

where $\omega^{(0)} \in \Omega^{(0)}$ is an element of the underlying probability space. We call $\{Z_t\}_{t \geq 0}$ the “*estimable latent process*” because we can indeed estimate it from the actual observations through, for example, pre-averaging [Podolskij and Vetter, 2009, Jacod et al., 2009, 2010, Mykland and Zhang, 2015b]. Additionally, assume the process $\{Z_t\}_{t \geq 0}$ is an Itô semimartingale, for example, if we assume $Z_t = f(X_t)$ [Li and Mykland, 2007] and $f(\cdot) \in \mathcal{C}^2$ so that $f(X_t)$ is also an Itô semimartingale³. Then we can also define another form of noise process, namely $\{\epsilon_t\}_{t \geq 0}$, which is not defined as the difference

²Although the noise is immaterial outside the observation times, it is not harm to assume there exist such a noise process in continuous time.

³The definition (3) suggests the possibility of our inability to recover the latent process $\{X_t\}$ from the noisy observations $\{Y_t\}$, since Z_t does not necessarily equal to X_t . More strikingly, as later discussed, this allows the correlation between the microstructure noise and the latent process.

between the observed process $\{Y_t\}_{t \geq 0}$ and the latent process $\{X_t\}_{t \geq 0}$ as tradition, but instead, it is defined theoretically via:

$$\epsilon_t \equiv Y_t - Z_t \quad (4)$$

we call $\{\epsilon_t\}_{t \geq 0}$ the “*distinguishable noise*”, which can be disentangled from the estimable latent process $\{Z_t\}_{t \geq 0}$ [Bandi and Russell, 2006].

Thirdly, we have a Markov kernel to provide a connection between the processes $\{X_t\}_{t \geq 0}$ and $\{Y_t\}_{t \geq 0}$ on different probability spaces: $Q_t(\omega^{(0)}, dy) : (\Omega^{(0)}, \mathcal{F}^{(0)}) \mapsto \mathbb{R}$, i.e., conditional on the whole latent process X , there exist a probability measure on the space $(\Omega^{(1)}, \mathcal{F}^{(1)})$ on which the observable process is defined⁴.

Thus, the latent Itô semimartingale $\{X_t\}_{t \geq 0}$, the observable process $\{Y_t\}_{t \geq 0}$, and the microstructure noise process $\{e_t\}_{t \geq 0}$ can be defined on the extended filter probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$:

$$\begin{cases} \Omega \equiv \Omega^{(0)} \times \Omega^{(1)}, \mathcal{F} \equiv \mathcal{F}^{(0)} \otimes \mathcal{F}^{(1)} \\ \mathcal{F}_t \equiv \bigcap_{s > t} \mathcal{F}_s^{(0)} \otimes \mathcal{F}_s^{(1)} \\ \mathbb{P}(d\omega^{(0)}, d\omega^{(1)}) \equiv \mathbb{P}^{(0)}(d\omega^{(0)}) \cdot \otimes_{t \geq 0} Q_t(\omega^{(0)}, dy_t) \end{cases} \quad (5)$$

The model setup above characterizes the underlying process and the general microstructure noise, which is one of the true state of nature we are interested in, but some features we can not directly observe. We need to find some good estimators and sound tests to make inference based on those observations.

2.2 Observation notation and assumptions

Suppose we focus on a compact interval $[0, T]$ on which ultra-high frequency data was observed. Let \mathcal{G} be the finest time grid, by which we can get the samples from the existing highest frequency sampling. Suppose we have $n + 1$ observation times, which are denoted by:

$$\mathcal{G} = \{t_0, t_1, t_2, \dots, t_n\}$$

Of course, we can do sparse sampling and only use partial observation data, for example take one sample from every K observations:

$$\mathcal{G}^{(k)} = \left\{ t_K, t_{k+K}, t_{k+2K}, \dots, t_{k+(\lfloor \frac{n}{K} \rfloor - 1) \cdot K} \right\}, \text{ where } k = 0, 1, 2, \dots, K - 1$$

However, in order to achieve identifiability and estimability, we have to make the following *identification assumption*:

$$dZ_t \equiv dX_t = b_t dt + \sigma_t dW_t \quad (6)$$

⁴This model combines the features of the two models considered in Li and Mykland [2007] and Jacod and Protter [2012] (or Jacod et al. [2009, 2010]). But it is not exactly the same as those models in the literature listed above, since we define another noise $\{\epsilon_t\}_{t \geq 0}$, and ϵ_{t_i} is not defined as the difference between the observations Y_{t_i} and the latent process X_{t_i} for the observation indexed by i , we define the noise ϵ_{t_i} through the difference between the observation Y_{t_i} and the value we can actually recover from the observation.

otherwise all the estimation methods will break down [Jacod et al., 2009]. And note that under the identification assumption (6), $\{e_t\}_{t \geq 0}$ and $\{\epsilon_t\}_{t \geq 0}$ are identical, and there is no correlation between noise and the latent process.

What's more, through the whole article, we assume:

- (i) Conditional on the latent variable(s), the noises occurred at different times are independent, i.e., $\epsilon_i \perp \epsilon_j$. This assumption simplifies the proof substantially.
- (ii) Define $g_t(\omega^{(0)}) = \int_{\mathbb{R}} (y_t - Z_t(\omega^{(0)}))^2 Q_t(\omega^{(0)}, dy_t)$, i.e., $g_{t_i}(\omega^{(0)}) = E(\epsilon_{t_i}^2 | \omega^{(0)})$, where $\omega^{(0)} \in \Omega^{(0)}$. By this definition, $g_t(\omega^{(0)})$ is also a stochastic process.
Note that $g_t(\omega^{(0)})$ could depend on more than one latent random variables, i.e., it is possible that $g_t(\omega^{(0)}) = g_t(X_t, Z_t, \sigma_t^2, \dots)$, in this case, we assume that $g_t(\cdot)$ is continuous;
- (iii) Mesh of the grid \mathcal{G} goes to zero, more specifically, $\max_i \Delta t_i = O(\frac{1}{n})$.
- (iv) $\forall l > 0, \exists M_{(5+2\delta_0, l)}$, s.t. $E(|\epsilon_{t_i}|^{5+2\delta_0} | \omega^{(0)}) \leq M_{(5+2\delta_0, l)}$, when $X_{t_i}, \sigma_{t_i}^2 \in [-l, l]$.

3 The non-stationarity problem and its remedy

3.1 Two-time scale estimator

The two-time scale realized volatility estimator (TSRV) [Zhang et al., 2005] is the first consistent estimator of integrated volatility $\int_0^T \sigma_t^2 dt$ using noisy high frequency financial data. It is defined as follows:

$$\widehat{\langle X, X \rangle}_T^{(TSRV, K)} \equiv [Y, Y]_T^{(avg, K)} - \frac{n - K + 1}{nK} [Y, Y]_T^{(all)} \quad (7)$$

where

$$\begin{aligned} [Y, Y]_T^{(all)} &\equiv [Y, Y]_{\mathcal{G}} = \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^2 \\ [Y, Y]_T^{(avg, K)} &= \frac{1}{K} \sum_{k=0}^{K-1} [Y, Y]_T^{(K, k)} \\ [Y, Y]_T^{(K, k)} &= \sum_{t_i \in \mathcal{G}^{(k)}} (Y_{t_i} - Y_{t_{i,-}})^2 \\ \mathcal{G}^{(k)} &= \{t_{k+K}, t_{k+2K}, \dots, t_{k+(\lfloor \frac{n}{K} \rfloor - 1) \cdot K}\}, \text{ for } k = 0, \dots, K-1 \\ &\text{and } t_{i,-} \text{ is the previous time point beside } t_i \text{ in grid } \mathcal{G}^{(k)} \end{aligned}$$

3.2 Edge effect under non-stationarity

In this section, we will focus on the question of quadratic variation estimation using high-frequency data contaminated by (possibly non-stationary) market microstructure noise.

In analogy with [Zhang et al. \[2005\]](#), we define:

$$M_T^{(1)} \equiv \frac{1}{\sqrt{n}} \sum_{i=0}^n (\epsilon_{t_i}^2 - g_{t_i}(\omega^{(0)})) \quad (8)$$

$$M_T^{(2)} \equiv \frac{1}{\sqrt{n}} \sum_{i=1}^n \epsilon_{t_i} \epsilon_{t_{i-1}} \quad (9)$$

$$M_T^{(3)} \equiv \frac{1}{\sqrt{n}} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_{i,-}} \quad (10)$$

where $\epsilon_{t_{i,-}}$ denotes the previous element in $\mathcal{G}^{(k)}$ when $t_i \in \mathcal{G}^{(k)}$, and $\epsilon_{t_{i,-}} = 0$ for $t_i = \min \mathcal{G}^{(k)}$. Note that $M_T^{(1)}$, $M_T^{(2)}$ and $M_T^{(3)}$ are the end-points of martingales with respect to filtration $\mathcal{F}_i = \sigma(\epsilon_{t_j}, j \leq i; X_t, \forall t)$. And we need a lemma⁵:

Lemma 1. *Under the model (1), (3) and (4), and the assumption (6), we have:*

$$[Y, Y]_T^{(all)} = [\epsilon, \epsilon]_T^{(all)} + O_p(1) \quad (11)$$

$$[Y, Y]_T^{(avg, K)} = [\epsilon, \epsilon]_T^{(avg, K)} + [Z, Z]_T^{(avg, K)} + O_p\left(\frac{1}{\sqrt{K}}\right) \quad (12)$$

Here we introduce a smaller summation: “ $\sum_{t_i \in \tilde{\mathcal{G}}^{(k)}}$ ”, which means the summation over the subset $\{\min \mathcal{G}^{(k)} + 1, \min \mathcal{G}^{(k)} + 1, \dots, \max \mathcal{G}^{(k)} - 1\}$, then we have the following finite-sample property:

$$\begin{aligned} & \widehat{\langle X, X \rangle}_T^{(TSRV, K)} - [Z, Z]_T^{(avg, K)} \\ &= 2 \cdot \underbrace{\frac{\sqrt{n}}{K} \left(\underbrace{\frac{K-1}{n}}_{o_p(1)} \cdot M_T^{(1)} + \underbrace{\frac{n-K+1}{n}}_{1+o_p(1)} \cdot M_T^{(2)} - M_T^{(3)} \right)}_{\text{Mixed Normal}} + \underbrace{O_p\left(\frac{1}{\sqrt{K}}\right)}_{\text{Negligible}} \\ & \quad + \underbrace{\frac{2(K-1)}{nK} \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} g_{t_i}(\omega^{(0)}) - \frac{n-2K+2}{nK} \left[\sum_{k=1}^K g_{\min \mathcal{G}^{(k)}}(\omega^{(0)}) + \sum_{k=1}^K g_{\max \mathcal{G}^{(k)}}(\omega^{(0)}) \right]}_{\text{Edge Effect in original TSRV}} \end{aligned} \quad (13)$$

which leads to the following lemma:

Lemma 2. *The finite-sample bias in the averaged realized variance using sparse samples due to noise (the different between $[Y, Y]_T^{(avg, K)}$ and $[Z, Z]_T^{(avg, K)}$) is:*

$$[Y, Y]_T^{(avg, K)} - [Z, Z]_T^{(avg, K)} = \underbrace{\sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} \frac{2}{K} g_{t_i}(\omega^{(0)}) + \sum_{k=1}^K \frac{1}{K} \left(g_{\min \mathcal{G}^{(k)}}(\omega^{(0)}) + g_{\max \mathcal{G}^{(k)}}(\omega^{(0)}) \right)}_{\text{bias in } [Y, Y]_T^{(avg, K)} \text{ due to noise}} + o_p(1)$$

⁵The proof for this lemma is very similar to that in [Zhang et al. \[2005\]](#), [Li and Mykland \[2007\]](#).

From the **Lemma 2** above⁶, we can see the noise in each time point does not contribute “equally” to the bias of averaged realized variance $[Y, Y]_T^{(avg, K)}$. In the beginning and ending parts of the sample points, the conditional second moments of noises are multiplied by the factor $\frac{1}{K}$, in contrast, the conditional second moments of noises in the middle of the whole sample are multiplied by the factor $\frac{2}{K}$. The correction in the next subsection 3.3 and the first two tests in **Section 4** and **5** are motivated by the the truncation of data at the beginning and the end of the time interval $[0, T]$.

3.3 Kalnina and Linton’s device

In [Kalnina and Linton \[2008\]](#), a parametric model was introduced to incorporate the diurnal and endogenous measurement error:

$$\begin{aligned} dX_t &= \mu_t dt + \sigma_t dW_t \\ Y_{t_i} &= X_{t_i} + \epsilon_{t_i} \\ \epsilon_{t_i} &= u_{t_i} + v_{t_i} \\ u_{t_i} &= \delta \gamma_n (W_{t_i} - W_{t_{i-1}}) \\ v_{t_i} &= m(t_i) + n^{-\frac{\alpha}{2}} \omega(t_i) e_{t_i}, \alpha \in [0, 1/2) \end{aligned}$$

where $e \perp\!\!\!\perp X$, i.i.d., with zero mean.

To the best of our knowledge, [Kalnina and Linton \[2008\]](#) is the first study which considered the border effect in TSRV due to the non-stationary microstructure noise, and they put forward a modified TSRV defined by:

$$[Y, Y]_T^{(avg, K)} - \frac{n - K + 1}{nK} [Y, Y]^{n\}$$

where $[Y, Y]^{n\} = \frac{1}{2} \left(\sum_{i=1}^{n-K} (Y_{t_{i+1}} - Y_{t_i})^2 + \sum_{i=K}^{n-1} (Y_{t_{i+1}} - Y_{t_i})^2 \right)$.

In next subsection, we will use this design to attack the non-stationarity problem under the general hidden Itô semimartingale model given in **Section 2**.

3.4 Sample-weighted TSRV

In this paper, we call the new TSRV using the modified version of realized variance in [Kalnina and Linton \[2008\]](#) as “sample-weighted TSRV”, which is defined as

$$\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} = [Y, Y]_T^{(avg, K)} - \frac{1}{K} [Y, Y]_T^{n\}$$

The sample-weighted TSRV enjoys the following asymptotic property under the general model in **Section 2**:

⁶The derivation of the finite sample property and the border effect can be found in the appendix 11.1 and 11.2.

Theorem 1. When we take $K = cn^{2/3}$ (the best possible order of TSRV), under the model assumptions (both on the latent process and the noise), then as $n \rightarrow \infty$,

$$n^{1/6} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \langle Z, Z \rangle_T \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \frac{8}{Tc^2} \int_0^T (g_t(\omega^{(0)}))^2 dt + c\xi^2 T \right) \quad (14)$$

where $\xi^2 = \frac{4}{3} \int_0^T \sigma_t^4 dt$.

The theorem tells us the sample-weighted TSRV in non-stationary noise setting enjoys the same asymptotic property as those of traditional TSRV in stationary noise setting [Zhang et al., 2005, Li and Mykland, 2007], in that the asymptotic distribution as well as the convergence rate remain unchanged, in other word, the asymptotic property of the sample-weighted TSRV is invariant with respect to non-stationary market microstructure noise.

4 Testing stationarity/non-stationarity: the first test

Based on the discussion in the previous sections, a natural question arises: could we find a statistical test to tell whether the market microstructure noise is stationary or non-stationary in a given period through the edge effect, by using the original TSRV and the sample-weighted TSRV simultaneously and comparing the estimates?

Consider testing the null hypothesis that the market microstructure noise is stationary:

$$H_0 : \epsilon_t \text{ is stationary} \longleftrightarrow H_1 : \epsilon_t \text{ is non-stationary}$$

assuming H_0 is true, both of the asymptotic distributions of the original TSRV and the sample-weighted TSRV are mixed normals. So, the asymptotic distribution of difference between the two different versions (after proper scaling) is also a mixed normal. Therefore, we can test the null $H_0 : \epsilon_t$ is stationary.

4.1 The first test $N(Y, K)_T^n$

Under stationary noise assumption, according to (53), the original TSRV estimator behaves like:

$$\widehat{\langle X, X \rangle}_T^{(TSRV, K)} - [Z, Z]_T^{(avg, K)} = \frac{2\sqrt{n}}{K} \left(\frac{K-1}{n} M_T^{(1)} + \frac{n-K+1}{n} M_T^{(2)} - M_T^{(3)} \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \quad (15)$$

When stationary noise assumption holds, the behavior of the sample-weighted TSRV is:

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - [Z, Z]_T^{(avg, K)} &= \frac{2\sqrt{n}}{K} \left(M_T^{(2)} - M_T^{(3)} \right) \\ &+ \frac{1}{\sqrt{K}} \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \end{aligned} \quad (16)$$

where,

$$\begin{aligned}
\underline{m}_T^{(1)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right] \\
\underline{m}_T^{(2)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \epsilon_{\mathcal{G}_{k+1}^{(\min)}} \epsilon_{\mathcal{G}_k^{(\min)}} \\
\bar{m}_T^{(1)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\max)}}^2 - g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) \right] \\
\bar{m}_T^{(2)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\max)}} \epsilon_{\mathcal{G}_{k-1}^{(\max)}}
\end{aligned}$$

which are also defined in the proof of **Theorem 1**.

From (50) in Section 11.1, we can notice that the error term $O_p\left(\frac{1}{\sqrt{K}}\right)$ in (15) and (16) comes from $-\frac{1}{K}R_2$ and $[Z, \epsilon]_T^{(avg, K)}$, which ultimately come from $[Y, Y]_T^{(avg, K)}$. So the difference between the two different versions of TSRV is due to the difference between $\frac{\bar{n}}{n}[Y, Y]_T^{(all)}$ and $\frac{1}{K}[Y, Y]_T^{\{n\}}$, which is of order $O_p\left(\frac{1}{K}\right)$:

$$\begin{aligned}
&\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \\
&= \underbrace{\frac{2(K-1)}{K\sqrt{n}} \left(M_T^{(2)} - M_T^{(1)} \right)}_{O_p\left(\frac{1}{\sqrt{n}}\right)} + \underbrace{\frac{1}{\sqrt{K}} \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right)}_{O_p\left(\frac{1}{\sqrt{K}}\right)} + O_p\left(\frac{1}{K}\right) \quad (17)
\end{aligned}$$

We can design our first test statistic $N(Y, K)_T^n$ defined by:

$$N(Y, K)_T^n \equiv \sqrt{K} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \right) \quad (18)$$

Remark Under the null hypothesis we have:

$$N(Y, K)_T^n = \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right) + o_p(1)$$

Our first test statistic has the following asymptotic property:

Theorem 2. If the noise process is stationary, under the assumptions of our model (1), (3) and (4) in Section 2,

$$N(Y, K)_T^n \xrightarrow{\mathcal{L}_s} \mathcal{MN}\left(0, 2E(\epsilon^4 | \omega^{(0)})\right) \quad (19)$$

Remark 1 We now investigate the behavior of our first test statistic under the alternative hypothesis (microstructure noise is not stationary).

By the previous calculation of the edge effect in the averaged realized variance $[Y, Y]_T^{avg, K}$, we know:

$$\begin{aligned} N(Y, K)_T &= \frac{2(K-1)}{\sqrt{nK}} \left(M_T^{(2)} - M_T^{(1)} \right) + \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right) + o_p(1) \\ &+ \frac{n-2K+2}{n\sqrt{K}} \left[\sum_{k=1}^K g_{\min \mathcal{G}^{(k)}}(\omega^{(0)}) + \sum_{k=1}^K g_{\max \mathcal{G}^{(k)}}(\omega^{(0)}) \right] - \frac{2(K-1)}{n\sqrt{K}} \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} g_{t_i}(\omega^{(0)}) \end{aligned} \quad (20)$$

So, we have:

$$\begin{aligned} N(Y, K)_T &= \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right) + o_p(1) \\ &+ \underbrace{\sqrt{K} \left[\overline{E(\epsilon^2 | \omega^{(0)})}^{(start)} + \overline{E(\epsilon^2 | \omega^{(0)})}^{(end)} - 2\overline{E(\epsilon^2 | \omega^{(0)})}^{(middle)} \right]}_{O_p(\sqrt{K})} \end{aligned} \quad (21)$$

where

$$\begin{aligned} \overline{E(\epsilon^2 | X)}^{(start)} &= \frac{1}{K} \sum_{k=1}^K g_{\min \mathcal{G}^{(k)}}(\omega^{(0)}) \\ \overline{E(\epsilon^2 | X)}^{(end)} &= \frac{1}{K} \sum_{k=1}^K g_{\max \mathcal{G}^{(k)}}(\omega^{(0)}) \\ \overline{E(\epsilon^2 | X)}^{(middle)} &= \frac{1}{n+1-2K} \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} g_{t_i}(\omega^{(0)}) \end{aligned} \quad (22)$$

Since $K = O_p\left(n^{\frac{2}{3}}\right)$ in our setup, this test statistic will explode when the noise is not stationary. Thus, the type-II error of this test is very small, and converges to zero as $K \rightarrow \infty$.

Remark 2 Actually, the calculation of the test statistic $N(Y, K)_T$ boils down to the calculation of the realized variances at the edges and the middle. Recall that

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(TSRV, K)} &= [Y, Y]_T^{(avg, K)} - \frac{n-K+1}{nK} [Y, Y]_T^{(all)} \\ \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} &= [Y, Y]_T^{(avg, K)} - \frac{1}{K} [Y, Y]_T^{\{n\}} \end{aligned}$$

so the difference between the two different versions of TSRV is:

$$\begin{aligned} &\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \\ &= \frac{n-K+1}{nK} [Y, Y]_T^{(all)} - \frac{1}{K} [Y, Y]_T^{\{n\}} \\ &= \frac{1}{2K} ([Y, Y]_{\mathcal{G}^{(\min)}} + [Y, Y]_{\mathcal{G}^{(\max)}}) - \frac{K-1}{nK} [Y, Y]_T^{(all)} \\ &= \frac{1}{2K} ([\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(\max)}}) - \frac{1}{n} [\epsilon, \epsilon]_{\mathcal{G}/(\mathcal{G}^{(\min)} \cup \mathcal{G}^{(\max)})} + O_p\left(\frac{1}{K}\right) \end{aligned} \quad (23)$$

For this reason, the test statistic $N(Y, K)_T^n$ can disclose the difference of the market microstructure noise level in the two edges of the mesh $\mathcal{G}^{(\min)}$, $\mathcal{G}^{(\max)}$ and the middle of the mesh $\mathcal{G}/(\mathcal{G}^{(\min)} \cup \mathcal{G}^{(\max)})$. We can show there are, in latter subsections, schemes which is able not only to reflect the heterogeneity in the edges and the middle, but also to capture the all (or almost all) of the information in the data. We will return to this in Section 5.1.

Remark 3 Based on **Theorem 2**, we know under the null

$$\frac{N(Y, K)_T^n}{\sqrt{2E(\epsilon^4|\omega^{(0)})}} = \frac{\sqrt{K} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \right)}{\sqrt{2E(\epsilon^4|\omega^{(0)})}} \xrightarrow{\mathcal{L}} N(0, 1)$$

So, when we know the market microstructure noises are mutually independent, we can use the quantity

$$\frac{N(Y, K)_T^n}{\sqrt{2} \cdot \sqrt{E(\epsilon^4|\omega^{(0)})}} = \frac{\sqrt{K} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \right)}{\sqrt{2} \cdot \sqrt{E(\epsilon^4|\omega^{(0)})}}$$

in our stationarity test.

4.2 Estimation of $E(\epsilon^4|\omega^{(0)})$

According to the **Lemma 3** below, when the noise is stationary, we can use the quarticity $[Y, Y, Y, Y]_T^{(all)} = \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^4$ computed using noisy data to estimate the 4-th moment of the noise $E(\epsilon^4|\omega^{(0)})$:

Lemma 3. If we define $h_t(\omega^{(0)}) \equiv E(\epsilon_t^4|\omega^{(0)})$, then under the assumption of our model, we have

$$\frac{1}{n} [Y, Y, Y, Y]_T^{(all)} = \frac{2}{T} \int_0^T h_t(\omega^{(0)}) dt + \frac{6}{T} \int_0^T g_t^2(\omega^{(0)}) dt + O_p\left(\frac{1}{\sqrt{n}}\right) \quad (24)$$

Remark If the noise is stationary, we have:

$$\frac{1}{2n} [Y, Y, Y, Y]_T^{(all)} | \omega^{(0)} \xrightarrow{\mathbb{P}} E(\epsilon^4|\omega^{(0)}) + 3 \left(E(\epsilon^2|\omega^{(0)}) \right)^2$$

So a natural estimate of $E(\epsilon^4|\omega^{(0)})$ is:

$$\begin{aligned} E(\widehat{\epsilon^4|\omega^{(0)}}) &= \frac{1}{2n} [Y, Y, Y, Y]_T^{(all)} - 3 \left(E(\widehat{\epsilon^2|\omega^{(0)}}) \right)^2 \\ &= \frac{1}{2n} \left([Y, Y, Y, Y]_T^{(all)} - \frac{3}{2n} ([Y, Y]_T^{(all)})^2 \right) \end{aligned}$$

Therefore we have $\frac{\sqrt{n} \cdot N(Y, K)_T^n}{\sqrt{[Y, Y, Y, Y]_T^{(all)} - \frac{3}{2n} ([Y, Y]_T^{(all)})^2}} =$

$$\frac{\sqrt{nK} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \right)}{\sqrt{[Y, Y, Y, Y]_T^{(all)} - \frac{3}{2n} ([Y, Y]_T^{(all)})^2}} \xrightarrow{\mathcal{L}} N(0, 1) \quad (25)$$

we use this result to test the stationarity of the market microstructure noise in section 9.2 (see Figure 7).

The estimator of $E(\epsilon^4 | \omega^{(0)})$ is not only used in the first test statistic but also used in the second test statistic in subsection 5.2.

5 Testing stationarity/non-stationarity: the second and third tests

5.1 The generic test

In order to effectively incorporate all the information about the noise stationarity contained in the data into our test, we designed the second test. In the following, we write K_n to indicate the tuning parameter is dependent on n .

Similar to the definition of the first test statistic, for a given subinterval $[t, t + s]$, define

$$N(Y, K_n)_{[t, t+s]} \equiv \sqrt{K_n} \left(\widehat{\langle X, X \rangle}_{[t, t+s]}^{(SW-TSRV, K_n)} - \widehat{\langle X, X \rangle}_{[t, t+s]}^{(TSRV, K_n)} \right) \quad (26)$$

where $\widehat{\langle X, X \rangle}_{[t, t+s]}^{(SW-TSRV, K_n)}$ and $\widehat{\langle X, X \rangle}_{[t, t+s]}^{(TSRV, K_n)}$ are estimators built upon the data in the time interval $[t, t + s]$.

Partition the fixed time interval $[0, \mathcal{T}]$ into subintervals $[T_i, T_{i+1}]$, where

$$0 = T_0 \leq T_1 \leq T_2 \leq \dots \leq T_r = \mathcal{T}$$

and each $[T_{i-1}, T_i]$ contains K_n observations. More, explicitly, we take $T_i = t_{iK_n}$. Then, we can use the moving window : $\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} |N(Y, K_n)_{[T_{i-1}, T_{i-1} + s_n]}|^2$ (we need square to avoid possible cancellation) with a suitably chosen window length s_n (in terms of subintervals), or something else to design the test. Generally, define:

$$V(Y, K_n, u)_{\mathcal{T}}^n \equiv \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} |N(Y, K_n)_{[T_{i-1}, T_{i-1} + s_n]}|^u \quad (27)$$

5.2 Test statistic $V(Y, K_n, 2)_{\mathcal{T}}^n$

Before the statement of the theorem, we need to introduce some notation: $r_n = \left\lfloor \frac{n}{K_n} \right\rfloor$.

For each $i = 1, 2, 3, \dots, r_n$, if we define

$$\begin{aligned} m_i^{(1)} &\equiv \frac{1}{\sqrt{K_n}} \sum_{k=1}^{K_n} \epsilon_{t_{(i-1)K_n+k}}^2 - g_{t_{(i-1)K_n+k}}(\omega^{(0)}) \\ m_i^{(2)} &\equiv \frac{1}{\sqrt{K_n}} \sum_{k=1}^{K_n} \epsilon_{t_{(i-1)K_n+k-1}} \epsilon_{t_{(i-1)K_n+k}} \\ m_i &\equiv m_i^{(1)} - m_i^{(2)} \end{aligned}$$

From (17) and (19), we know under the null hypothesis,

$$\begin{aligned} N(Y, K_n)_{[T_{i-1}, T_{i-1}+s_n]} &= m_i^{(1)} - m_i^{(2)} + m_{i+s_n}^{(1)} - m_{i+s_n}^{(2)} + o_p(1) \\ &= m_i + m_{i+s_n} + o_p(1) \end{aligned}$$

therefore each $N(Y, K_n)_{[T_{i-1}, T_{i-1}+s_n]}$ is asymptotically mixed normal, and we have the following result:

Theorem 3. ($V(Y, K_n, 2)_{\mathcal{T}}^n$ *under the null*) Under the model (1), (3) and (4), assume the noise process is stationary, and $K_n \rightarrow \infty$, $r_n \rightarrow \infty$, $r_n = o(n)$, $s_n \rightarrow \infty$, $s_n = o(\sqrt{r_n})$, then the test statistic

$$V(Y, K_n, 2)_{\mathcal{T}}^n = \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} |N(Y, K_n)_{[T_{i-1}, T_{i-1}+s_n]}|^2 \quad (28)$$

has the following asymptotic property:

$$\sqrt{r_n - s_n + 1} \left(V(Y, K_n, 2)_{\mathcal{T}}^n - 2E(\epsilon^4 | \omega^{(0)}) \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN}(0, \eta^2) \quad (29)$$

where $\eta = 2\sqrt{6} \cdot \sqrt{(E(\epsilon^4 | \omega^{(0)}))^2 - E(\epsilon^4 | \omega^{(0)}) (E(\epsilon^2 | \omega^{(0)}))^2 + (E(\epsilon^2 | \omega^{(0)}))^4}$.

Remark Based on **Theorem 3**, we have the following convergence result:

$$\frac{\sqrt{r_n - s_n + 1} \left(V(Y, K, 2)_{\mathcal{T}}^n - \frac{1}{n} \left([Y, Y, Y, Y]_{\mathcal{T}}^{(all)} - \frac{3}{2n} ([Y, Y]_{\mathcal{T}}^{(all)})^2 \right) \right)}{\sqrt{\hat{\eta}^2}} \xrightarrow{\mathcal{L}} N(0, 1)$$

where $\hat{\eta}^2$ is the plug-in estimate of η^2 :

$$\begin{aligned} \hat{\eta}^2 &= \frac{6}{n^2} \left([Y, Y, Y, Y]_{\mathcal{T}}^{(all)} - \frac{3}{2n} ([Y, Y]_{\mathcal{T}}^{(all)})^2 \right)^2 \\ &\quad - \frac{3}{n^3} \left([Y, Y, Y, Y]_{\mathcal{T}}^{(all)} - \frac{3}{2n} ([Y, Y]_{\mathcal{T}}^{(all)})^2 \right) \left([Y, Y]_{\mathcal{T}}^{(all)} \right)^2 + \frac{3}{2n^4} \left([Y, Y]_{\mathcal{T}}^{(all)} \right)^4 \end{aligned}$$

we use this result to test the stationarity of the market microstructure noise in section 9.2 (see Figure 8).

To prove **Theorem 3**, we need a additional lemma:

Lemma 4. *Under the null hypothesis that the microstructure noise is stationary, and under the moment assumptions on the noise process $\{\epsilon_t\}_{t \geq 0}$, we have the following relations for each $i \in \{1, 2, \dots, r_n\}$*

$$\begin{aligned} E(m_i^2 | \omega^{(0)}) &= E(\epsilon^4 | \omega^{(0)}) \\ E(m_i^4 | \omega^{(0)}) &= 6 \left[(E(\epsilon^4 | \omega^{(0)}))^2 - E(\epsilon^4 | \omega^{(0)}) (E(\epsilon^2 | \omega^{(0)}))^2 + (E(\epsilon^2 | \omega^{(0)}))^4 \right] + O_p \left(\frac{1}{K} \right) \end{aligned}$$

The proofs of **Lemma 4** and **Theorem 3** are provided in the appendices **11.6** and **11.7**.

We can also define another quantity $V'(Y, K_n, 2)_{\mathcal{T}}^n$ built upon $N(Y, K_n)$ over non-overlapping intervals:

$$V'(Y, K_n, 2)_{\mathcal{T}}^n \equiv \frac{1}{\lfloor r_n/s_n \rfloor} \sum_{j=1}^{\lfloor r_n/s_n \rfloor} \left| N(Y, K_n)_{[T_{(j-1)s_n}, T_{js_n}]} \right|^2$$

A corollary describing the asymptotic property of $V'(Y, K_n, 2)_{\mathcal{T}}^n$ follows directly from the **Theorem 3**:

Corollary 1. *Under the same conditions as in **Theorem 3**, $V'(Y, K_n, 2)_{\mathcal{T}}^n$ has the following asymptotic properties under the null hypothesis:*

$$\sqrt{\lfloor r_n/s_n \rfloor} (V'(Y, K_n, 2)_{\mathcal{T}}^n - 2E(\epsilon^4 | \omega^{(0)})) \xrightarrow{\mathcal{L}-s} \mathcal{MN}(0, \eta^2) \quad (30)$$

Remark It is a little bit surprising when we compare **Corollary 1** with **Theorem 3**, since the limiting mixed normals have the same asymptotic variance although the convergence rate of the former is lower. However, the results only demonstrate the limiting behaviors. $V'(Y, K_n, 2)_{\mathcal{T}}^n$ required less computation, while $V(Y, K_n, 2)_{\mathcal{T}}^n$ is more accurate in terms of asymptotic approximation because of its higher rate of convergence.

5.3 An equivalent test: the third test $U(Y, K_n, 2)_{\mathcal{T}}^n$

However, there is also an edge effect in the second test statistic (28) (coming from the first $s_n K_n$ and the last $s_n K_n$ observations). Motivated by the **Remark 2** of the first test statistic (18), we can design another test statistic with a similar asymptotic properties with $V(Y, K_n, 2)_{\mathcal{T}}^n$ under the null, but has a smaller edge effect:

$$U(Y, K_n, 2)_{\mathcal{T}}^n \equiv \frac{1}{r_n} \sum_{i=1}^{r_n-1} \frac{\left| [Y, Y]_{[T_i, T_{i+1}]}^{(all)} - [Y, Y]_{[T_{i-1}, T_i]}^{(all)} \right|^2}{4K_n} \quad (31)$$

Theorem 4. ($U(Y, K_n, 2)_T^n$ **under the null**) Under the model (1), (3) and (4), assume the noise process is stationary, suppose $K_n \rightarrow \infty$, $r_n \rightarrow \infty$ and $r_n = o(n)$, then the test statistic

$$U(Y, K_n, 2)_T^n = \frac{1}{r_n} \sum_{i=1}^{r_n-1} \frac{\left| [Y, Y]_{[T_i, T_{i+1}]}^{(all)} - [Y, Y]_{[T_{i-1}, T_i]}^{(all)} \right|^2}{4K_n}$$

has the following asymptotic property:

$$\sqrt{r_n} \left(U(Y, K_n, 2)_T^n - 2E(\epsilon^4 | \omega^{(0)}) \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN}(0, \eta^2) \quad (32)$$

thus we have the following convergence result:

$$\frac{\sqrt{r_n} \left(U(Y, K_n, 2)_T^n - \frac{1}{n} \left([Y, Y, Y, Y]_T^{(all)} - \frac{3}{2n} \left([Y, Y]_T^{(all)} \right)^2 \right) \right)}{\sqrt{\hat{\eta}^2}} \xrightarrow{\mathcal{L}} N(0, 1)$$

where η and $\hat{\eta}$ are the same as those in **Theorem 3**.

Remark 1 The testing theorems in this paper, **Theorem 2**, **Theorem 3**, **Corollary 1** and **Theorem 4** are robust to finitely many jumps.

These test statistics are built upon the realized variances at the two edges, which involve the realized variances of the fastest time scale on the edges. When jump component of finite activity exists in the process, there are three components in the realized variance:

- (1) the variation in the latent Itô process, which is of order $O_p(1)$;
- (2) the variation in the noise, which is of order $O_p(n)$;
- (3) the variation due to jumps, which is of order $O_p(1)$ because of its finite activities.

Formally, if we add a finite-activity jump process into the model (33):

$$X_t = X_0 + \underbrace{\int_0^t b_s ds + \int_0^t \sigma_s dW_s}_{X_t^c} + \underbrace{J_t}_{X_t^d} \quad (33)$$

where $\{J_t\}_{t \geq 0}$ is a pure jump process which only has finitely many activities over a fixed time interval. Because the noise $\{\epsilon_t\}_{t \geq 0}$ is independent of the latent process $\{X_t\}_{t \geq 0}$, by the similar argument in the proof of lemma 1 in [Li and Mykland \[2007\]](#)⁷, we have the result similar to that of **Lemma 1**:

$$[Y, Y]_T^{(all)} = [\epsilon, \epsilon]_T^{(all)} + \underbrace{[X^d, \epsilon]_T}_{O_p(1)} + \underbrace{\int_0^T \sigma_t^2 dt + \sum_{t \in (0, T]} (\Delta X_t^d)^2}_{O_p(1)} + o_p(1) \quad (34)$$

⁷Lemma 1 on p. 606 in [Li and Mykland \[2007\]](#)

which suggests that normalized realized variance of the fastest time scale $\frac{1}{2n}[Y, Y]_T^{(all)}$ is able to consistently estimates the quantity $E(\epsilon^2|\omega^{(0)})$ provided the noise is stationary even if there exist jumps with finite activities, i.e., (11) still holds. For this reason, the asymptotic distributions remain the same for the test statistics under the null.

Remark 2 The **Theorem 3** and **4** give the asymptotic distributions of $V(Y, K_n, 2)_T^n$ and $U(Y, K_n, 2)_T^n$ under the null hypothesis, which aid us to control the type-I error. In **Section 5**, we will study the asymptotic behaviors under the alternative hypothesis. Besides the type-II error, the behaviors under non-stationary hypothesis can offer us other insights such as aggregated non-stationarity level or accumulated changes in the noise variance. Since the market microstructure noise captures the liquidity level of the market, the statistic $V(Y, K_n, 2)_T^n$ and $U(Y, K_n, 2)_T^n$ can also manifest the (aggregated) liquidity changes.

6 The behaviors under non-stationary noise and aggregate liquidity risks

6.1 Quadratic variation of $g_t(\omega^{(0)})$: aggregate liquidity risk

In the case that we have an equi-distant sample grid over a long compact time interval $[0, \mathcal{T}]$, we can apply the technique provided in ? to design a device to measure the aggregated liquidity risk over a long period.

Similar to the observation scheme in the second test, suppose we partition the whole time interval into r_n disjoint subintervals $(T_{i-1}, T_i]$ for $i = 1, 2, \dots, r_n$, especially, note that $T_0 = 0$ and $T_{r_n} = \mathcal{T}$, and in each subinterval we have K_n observations. Furthermore, assume $\Delta T = T_i - T_{i-1}$, $\forall i = 1, 2, \dots, r_n$ (equivalently assume we adapt a regular observation scheme).

From (49), we know:

$$[\epsilon, \epsilon]_{(T_{i-1}, T_i]}^{(all)} = 2\sqrt{K_n} \left(M_{(T_{i-1}, T_i]}^{(1)} - M_{(T_{i-1}, T_i]}^{(2)} \right) + 2 \sum_{t_j \in (T_{i-1}, T_i]} g_{t_j}(\omega^{(0)}) + O_p(1)$$

where

$$\begin{aligned} M_{(T_{i-1}, T_i]}^{(1)} &\equiv \frac{1}{\sqrt{K_n}} \sum_{t_j \in (T_{i-1}, T_i]} \left(\epsilon_{t_j}^2 - g_{t_j}(\omega^{(0)}) \right) \\ M_{(T_{i-1}, T_i]}^{(2)} &\equiv \frac{1}{\sqrt{K_n}} \sum_{t_j \in (T_{i-1}, T_i]} \epsilon_{t_{j-1}} \epsilon_{t_j} \end{aligned}$$

which are asymptotically mixing normal. Since $[Y, Y]_{(T_{i-1}, T_i]}^{(all)} = [\epsilon, \epsilon]_{(T_{i-1}, T_i]}^{(all)} + O_p(1)$, we have:

$$\frac{\Delta T}{2K_n} [Y, Y]_{(T_{i-1}, T_i]}^{(all)} = \int_{T_{i-1}}^{T_i} g_t(\omega^{(0)}) dt + o_p(1) \quad (35)$$

In this section, we assume $g \in \mathcal{C}^2$. By Itô lemma, $\{g_t\}_{t \geq 0}$ is an Itô semi-martingale. Let $G_i \equiv \int_{T_{i-1}}^{T_i} g_t(\omega^{(0)}) dt$ and $\hat{G}_i \equiv \frac{\Delta T}{2K_n} [Y, Y]_{(T_{i-1}, T_i]}^{(all)}$, according to the “integral-to-spot device” in Mykland and Zhang [2015a]⁸, we know:

$$\frac{3}{2(\Delta T)^2} \sum_{i=1}^{r_n} (G_{i+1} - G_i)^2 \xrightarrow{\mathbb{P}} [g, g]_{\mathcal{T}-} \quad (36)$$

where $[\theta, \theta]_{\mathcal{T}-} = \lim_{t \nearrow \mathcal{T}} [\theta, \theta]_t$. Under additional some regularity conditions given in assumption 1 and theorem 2 in Mykland and Zhang [2015a]⁹, we have:

$$\frac{3}{2(\Delta T)^2} \sum_{i=1}^{r_n} (\hat{G}_{i+1} - \hat{G}_i)^2 \xrightarrow{\mathbb{P}} [g, g]_{\mathcal{T}-} + (\text{possibly additional terms}) \quad (37)$$

where $[g, g]_{\mathcal{T}-} = \lim_{t \nearrow \mathcal{T}} [g, g]_t$.

Since $[g, g]_{\mathcal{T}}$, the quadratic variation of $\{g_t\}_{t \geq 0}$ over $(0, \mathcal{T})$, which is a reasonable measure of the “aggregate” variation of the process $\{g_t\}_{t \geq 0}$, so we can interpret $[g, g]_{\mathcal{T}}$ as “aggregate liquidity risks” in the term of financial economics.

Note that $\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n = \frac{1}{(\Delta T)^2} \sum_{i=1}^{r_n-1} (\hat{G}_{i+1} - \hat{G}_i)^2$, thus, by studying the limiting distribution of $U(Y, K_n, 2)_{\mathcal{T}}^n$ under the alternative hypothesis, we can discover the possible additional terms and provide the central limit theorem for (37). Luckily, as it turns out, the additional terms in (37) is zero, we will see that in **Theorem 5** in the following subsection.

6.2 The behavior of $U(Y, K, 2)_{\mathcal{T}}^n$ in presence of non-stationary noises

Theorem 5. (*$U(Y, K_n, 2)_{\mathcal{T}}^n$ under the alternative*) Assume our model assumptions with regular sampling scheme, and adapt the same notation for r_n and K_n as in **Theorem 4** (the subscript n indicates the dependence of r_n and K_n on n), but with further assumptions that $\frac{r_n}{K_n} \rightarrow 0$ and $\frac{r_n^2}{K_n} \rightarrow \infty$. Furthermore, assume $g_t(\omega^{(0)}) = E(\epsilon_t^2 | \omega^{(0)})$ is

⁸The theorem 1 (“the integral-to-spot device”) in Mykland and Zhang [2015a]: for a semimartingale θ_t on $[0, \mathcal{T}]$, let $\Theta_{(T_i, T_{i+q}]} = \int_{T_i}^{T_{i+q}} \theta_t dt$, and $QV_q(\Theta) = \frac{1}{q} \sum_{i=q}^{r_n-q} (\theta_{(T_i, T_{i+q}]} - \theta_{(T_{i-q}, T_i]})^2$, then

$$\frac{1}{(q\Delta T)^2} QV_q(\Theta) \xrightarrow{\mathbb{P}} \frac{2}{3} [\theta, \theta]_{\mathcal{T}-}$$

as $q \rightarrow \infty$ and $q\Delta T \rightarrow 0$.

⁹Basically specking, the required assumption is the conditions for standard stable convergence plus additional restriction on edge effects, then

$$QV_q(\Theta) = \frac{2}{3} (q\Delta)^2 [\theta, \theta]_{\mathcal{T}-} + 2n^{-\alpha} [L, L]_{\mathcal{T}-} + o_p(n^{-\alpha})$$

where L_t is a limiting quantity in the stable convergence and it is a nonvanishing local martingale, α has something to do with the magnitude of the edge effect (whose magnitude is of the order $o_p(n^{-\alpha})$).

an Itô process (in time t), $d\langle g, g \rangle_t = (\sigma_t^{(g)})^2 dt$, and $(\sigma_t^{(g)})^2$ is also an Itô process and locally bounded, Then we have:

$$\sqrt{r_n} \left(\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n - \frac{2}{3} \langle g, g \rangle_{\mathcal{T}} - \frac{2r_n}{K_n \mathcal{T}} \int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \frac{2\mathcal{T}}{3} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt \right) \quad (38)$$

where $h_t(\omega^{(0)}) \equiv E(\epsilon_t^4 | \omega^{(0)})$.

Remark 1 Since $\frac{r_n}{K_n} \rightarrow 0$ as $n \rightarrow \infty$ and $\langle g, g \rangle_{\mathcal{T}}$ is finitely positive, the test statistic $U(Y, K_n, 2)_{\mathcal{T}}^n$ indeed explodes when the market microstructure noise process is non-stationary. So in term of in-fill asymptotics, $U(Y, K_n, 2)_{\mathcal{T}}^n$ is a powerful test statistic to discover non-stationarity in market microstructure noise.

Remark 2 Since $\frac{r_n}{K_n} \rightarrow 0$ as $n \rightarrow \infty$ and $\int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt$ is finitely positive, $\frac{3r_n}{2K_n} U(Y, K_n, 2)_{\mathcal{T}}^n$ is a consistent estimator of $\langle g, g \rangle_{\mathcal{T}}$, i.e., there is no additional term in (37). However, we can rewrite (38) as following form:

$$\sqrt{r_n} \left(\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n - \frac{2}{3} \langle g, g \rangle_{\mathcal{T}} \right) - \frac{2r_n^{3/2}}{K_n \mathcal{T}} \int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \frac{2\mathcal{T}}{3} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt \right) \quad (39)$$

depending the relation between the number of blocks and number of observations within each block, we have three different situations:

- (1) if $K_n = o_p(r_n^{3/2})$, i.e., $\frac{r_n^{3/2}}{K_n} \rightarrow \infty$, $\sqrt{r_n} \left(\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n - \frac{2}{3} \langle g, g \rangle_{\mathcal{T}} \right)$ converges to a mixing normal $\mathcal{MN} \left(0, \frac{2\mathcal{T}}{3} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt \right)$ plus an diverging bias $\frac{2r_n^{3/2}}{K_n \mathcal{T}} \int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt \rightarrow \infty$;
- (2) if $K_n = O_p(r_n^{3/2})$, then we know $\sqrt{r_n} \left(\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n - \frac{2}{3} \langle g, g \rangle_{\mathcal{T}} \right)$ converges to a non-zero mixing normal $\mathcal{MN} \left(\frac{2c}{\mathcal{T}} \int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt, \frac{2\mathcal{T}}{3} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt \right)$, where c is a finite constant;
- (3) if $r_n = o_p(K_n^{2/3})$, i.e., $\frac{r_n^{3/2}}{K_n} \rightarrow 0$ as $n \rightarrow \infty$, the term $\frac{2r_n^{3/2}}{K_n \mathcal{T}} \int_0^{\mathcal{T}} h_t(\omega^{(0)}) dt$ is negligible and $\sqrt{r_n} \left(\frac{r_n}{K_n} U(Y, K_n, 2)_{\mathcal{T}}^n - \frac{2}{3} \langle g, g \rangle_{\mathcal{T}} \right)$ converges to a mixing normal $\mathcal{MN} \left(0, \frac{2\mathcal{T}}{3} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt \right)$.

Remark 3 Careful reader who pay attention to the proof of **Theorem 5** will notice that the error term (E2) in (76) also contributes to the asymptotic variance in the limit distribution (38), although its contribution is negligible in the asymptotic setting. However, for the sake of finite-sample performance, for example, to get a more accurate confidence interval for the aggregate liquidity risk, we suggest to use the estimate of

$$\underbrace{\frac{3\mathcal{T}}{2} \int_0^{\mathcal{T}} (\sigma_t^{(g)})^4 dt}_{\text{due to discretization (non-vanishing)}} + \underbrace{\frac{24r_n^2}{\mathcal{T}K_n^2} \int_0^{\mathcal{T}} [h_t^2(\omega^{(0)}) - h_t(\omega^{(0)})g_t^2(\omega^{(0)}) + g_t^4(\omega^{(0)})] dt}_{\text{due to market microstructure noise (vanishing)}}$$

as the estimate of asymptotic approximation to the finite-sample variance, in order to avoid the situation in which we underestimate the finite-sample variance and become overoptimistic about the accuracy of our estimate. Although it is worthwhile to find a good estimate for the “finite-sample” variance, it is beyond the scope of the discussion of this paper.

Another reason to study the behavior of $U(Y, K_n, 2)_T^n$ under the alternative hypothesis is that it also provides the CLT for the estimation of aggregate liquidity risk in 6.1, which is given by the following corollary following directly from the **Theorem 5**:

Corollary 2. *Under the same conditions as in **Theorem 5**, and suppose $r_n = o_p(K_n^{2/3})$, then we know:*

$$\sqrt{r_n} \left(\frac{3}{2(\Delta T)^2} \sum_{i=1}^{r_n} (\hat{G}_{i+1} - \hat{G}_i)^2 - \langle g, g \rangle_T \right) \xrightarrow{\mathcal{L}} \mathcal{MN} \left(0, \frac{3T}{2} \int_0^T (\sigma_t^{(g)})^4 dt \right) \quad (40)$$

Thus, we can consistently estimate $\langle g, g \rangle_T$ by $\frac{3}{2(\Delta T)^2} \sum_{i=1}^{r_n} (\hat{G}_{i+1} - \hat{G}_i)^2$, where $\hat{G}_i \equiv \frac{\Delta T}{2K_n} [Y, Y]_{(T_{i-1}, T_i]}^{(all)}$, and the rate of convergence is $\sqrt{r_n}$.

6.3 The behavior of $V(Y, K, 2)_T^n$ in the presence of non-stationary noise

How $V(Y, K, 2)_T^n$ behaves when the noise is not-stationary is quite important (in this case at hand, $g(\cdot)$ is not constant, instead it is a function of time as well as possible latent variable(s)). For instance, if the test statistic $V(Y, K, 2)_T^n$ tends to be large when the market microstructure noise is non-stationary, then the test statistic can easily detect non-stationary market microstructure and reject the null hypothesis when alternative hypothesis holds.

Below, we use a heuristic argument to look at the behavior of the test statistic under non-stationary noise.

By (21), we know for a small time interval $[T_i - \Delta T, T_i + \Delta T]$ which contains large amount of observations, namely K observations, we have:

$$\begin{aligned} N(Y, K)_{[T_i - \Delta T, T_i + \Delta T]} &\approx \text{Asym. Gaussian Martingales} \\ &+ \sqrt{K} \left[g_{T_i - \Delta T}(\omega^{(0)}) - \frac{1}{2\Delta T} \int_{T_i - \Delta T}^{T_i + \Delta T} g_t(\omega^{(0)}) dt + g_{T_i + \Delta T}(\omega^{(0)}) \right] \end{aligned}$$

we assume $g_t(\cdot)$ is a function of d latent variable(s), denoted by $\{f_t^{(k)}(\omega^{(0)})\}_{t \geq 0}$, $k = 1, \dots, d$, and assume $g_t(\cdot) \in \mathcal{C}^2(\mathbb{R}^d)$ with nonzero second-order derivative for each $t \in [0, T]$ and $g_t(\cdot)$ is differentiable w.r.t. time variable. Besides, assume that for each k , $\{f_t^{(k)}(\omega^{(0)})\}_{t \geq 0}$ is also an Itô process with volatility $\{\sigma_t^{(k)}\}_{t \geq 0}$, and the spot correlation between $f_t^{(j)}(\omega^{(0)})$ and $f_{t+s}^{(k)}(\omega^{(0)})$ is $\rho_t^{(jk)} \mathbf{1}_{\{s=0\}}$. In other words, we can write

$$g_t(\omega^{(0)}) \equiv g \left(t, \omega^{(0)}, f_t^{(1)}(\omega^{(0)}), f_t^{(2)}(\omega^{(0)}), \dots, f_t^{(d)}(\omega^{(0)}) \right) \quad (41)$$

also note that

$$g_{T_i-\Delta T}(\omega^{(0)}) - 2g_{T_i}(\omega^{(0)}) + g_{T_i+\Delta T}(\omega^{(0)}) = (g_{T_i+\Delta T}(\omega^{(0)}) - g_{T_i}(\omega^{(0)})) - (g_{T_i}(\omega^{(0)}) - g_{T_i-\Delta T}(\omega^{(0)}))$$

When $\Delta T \rightarrow 0$, by Itô formula:

$$\begin{aligned} g_{T_i+\Delta T}(\omega^{(0)}) - g_{T_i}(\omega^{(0)}) &\approx \frac{\partial}{\partial t} g_{T_i}(\omega^{(0)}) \Delta T + \sum_{k=1}^d \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \Delta f_{T_i+\Delta T}^{(k)}(\omega^{(0)}) \\ &\quad + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \frac{\partial^2 g_{T_i}(\omega^{(0)})}{\partial f_t^{(j)}(\omega^{(0)}) \partial f_t^{(k)}(\omega^{(0)})} \rho_{T_i}^{(jk)} \sigma_{T_i}^{(j)} \sigma_{T_i}^{(k)} \Delta T \\ g_{T_i}(\omega^{(0)}) - g_{T_i-\Delta T}(\omega^{(0)}) &\approx \frac{\partial}{\partial t} g_{T_i}(\omega^{(0)}) \Delta T + \sum_{k=1}^d \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \Delta f_{T_i}^{(k)}(\omega^{(0)}) \\ &\quad + \frac{1}{2} \sum_{j=1}^d \sum_{k=1}^d \frac{\partial^2 g_{T_i}(\omega^{(0)})}{\partial f_t^{(j)}(\omega^{(0)}) \partial f_t^{(k)}(\omega^{(0)})} \rho_{T_i}^{(jk)} \sigma_{T_i}^{(j)} \sigma_{T_i}^{(k)} \Delta T \end{aligned}$$

where $\Delta f_{T_i+\Delta T}^{(k)}(\omega^{(0)}) = f_{T_i+\Delta T}^{(k)}(\omega^{(0)}) - f_{T_i}^{(k)}(\omega^{(0)})$ and $\Delta f_{T_i}^{(k)}(\omega^{(0)}) = f_{T_i}^{(k)}(\omega^{(0)}) - f_{T_i-\Delta T}^{(k)}(\omega^{(0)})$, so

$$g_{T_i-\Delta T}(\omega^{(0)}) - 2g_{T_i}(\omega^{(0)}) + g_{T_i+\Delta T}(\omega^{(0)}) \approx \sum_{k=1}^d \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \left(\Delta f_{T_i+\Delta T}^{(k)}(\omega^{(0)}) - \Delta f_{T_i}^{(k)}(\omega^{(0)}) \right)$$

$$\left[g_{T_i-\Delta T}(\omega^{(0)}) - 2g_{T_i}(\omega^{(0)}) + g_{T_i+\Delta T}(\omega^{(0)}) \right]^2 \approx (\text{mean-0 martingale}) +$$

$$\begin{aligned} &\sum_{j=1}^d \sum_{k=1}^d \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(j)}(\omega^{(0)})} \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \left(\Delta f_{T_i}^{(j)}(\omega^{(0)}) \cdot \Delta f_{T_i}^{(k)}(\omega^{(0)}) \right) \\ &+ \sum_{j=1}^d \sum_{k=1}^d \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(j)}(\omega^{(0)})} \frac{\partial g_{T_i}(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \left(\Delta f_{T_i+\Delta T}^{(j)}(\omega^{(0)}) \cdot \Delta f_{T_i+\Delta T}^{(k)}(\omega^{(0)}) \right) \end{aligned}$$

Therefore, under the alternative our test statistic $V(Y, K, 2)_{\mathcal{T}}^n$ behaves like

$$\begin{aligned} V(Y, K, 2)_{\mathcal{T}}^n &= \frac{1}{r_n - s_n + 1} \sum_{T_i} \left| N(Y, K)_{[T_i-\Delta T, T_i+\Delta T]} \right|^2 \\ &\approx (\text{Martingales}) + (\text{average of Asym. Chi-squares}) \\ &\quad + \frac{2K}{\mathcal{T}} \left[\sum_{j=1}^d \sum_{k=1}^d \int_0^T \frac{\partial g_t(\omega^{(0)})}{\partial f_t^{(j)}(\omega^{(0)})} \frac{\partial g_t(\omega^{(0)})}{\partial f_t^{(k)}(\omega^{(0)})} \rho_t^{(jk)} \sigma_t^{(j)} \sigma_t^{(k)} dt \right] \end{aligned}$$

Based on this finding, the second test statistic will explode under non-stationary microstructure noise, and in this situation, this test can easily distinguish non-stationary noise.

7 Noise functional dependency and model extension

In our model (1), (3) and (4), there is a conditional structure posted on the market microstructure noise $\{\epsilon_t\}_{t \geq 0}$, in other words, we represented the microstructure noise via a Markov kernel $Q_t(\omega^{(0)}, d\omega^{(1)})$ for each time t , and we denoted the conditional second moment of the noise by $g_t(\omega^{(0)}) = E(\epsilon_t^2 | \omega^{(0)})$, which is a random function on the probability space $(\Omega^{(0)}, \mathcal{F}^{(0)}, \mathbb{P}^{(0)})$.

Generally, the random function $g_t(\omega^{(0)})$ could depend on various latent variables, and the form of $Q_t(\omega^{(0)}, d\omega^{(1)})$ allows a wide range of correlation structures between the efficient price process $\{X_t\}_{t \geq 0}$ and the market microstructure noise $\{\epsilon_t\}_{t \geq 0}$. In this section, we bring further discussion on functional dependency of noise variance and related model extension, namely an elementary inference theory on $g_t(\omega^{(0)})$ and the implication of abandoning the identification assumption $Z_t = X_t, \forall t \in [0, T]$ from Section 2.

7.1 Regression: market microstructure noises and spot volatilities

We will conduct time series linear regression of $g_t(\omega^{(0)})$ on various latent variables, for example, $\sigma_t^2(\omega^{(0)})$.

We can use the TSRV to estimate σ_t^2 's using the samples in a narrow window by sample-weighted TSRV, and estimate local noise levels using the same samples by realized variance of the fastest time scale, i.e., $\widehat{E\epsilon^2} = \frac{1}{2n} [Y, Y]_\Lambda^{(all)}$, $\hat{\sigma}^2 = \frac{1}{|\Lambda|} \widehat{\langle X, X \rangle}_\Lambda^{(SW-TSRV)}$, Λ is some small time interval (e.g. a few hours in a trading day), and $|\Lambda|$ is its length.

In this subsection, we assume (at least locally) that the latent market microstructure noise variance and the latent volatility are correlated in the following manner:

$$E\epsilon_t^2 = \beta\sigma_t^2 + \alpha + \zeta_t \quad (42)$$

where ζ_t is the component of the microstructure noise variance at time t which is attributed to variables other than the spot volatility σ_t^2 . Then we can conduct linear regression on these pairs of volatility-noise estimates $(\hat{\sigma}_t^2, \widehat{E\epsilon_t^2})$:

$$\widehat{E\epsilon_t^2} = \hat{\beta}_n \hat{\sigma}_t^2 + \hat{\alpha}_n + \eta_t^{(n)} \quad (43)$$

where n is the number of observation in the small time interval Λ , and $\eta_t^{(n)}$ denotes a component in the noise variance not captured by the volatility estimator $\hat{\sigma}_t^2$, which is independent of $\hat{\sigma}_t^2$, β_n , α_n . Beside, we use n in the subscripts of estimators $\hat{\alpha}_n$ and $\hat{\beta}_n$ to emphasize that the values of the estimators in (43) depend on the sample size n , and the distribution of $\eta_t^{(n)}$ also depends on n .

Lemma 5. Suppose (42) holds, then the coefficient estimates $\hat{\beta}_n$ and $\hat{\alpha}_n$ in the linear regression (43) between spot estimates $\widehat{E\epsilon_t^2}$ and $\hat{\sigma}_t^2$ converge to the corresponding values β and α in (42), i.e., the linear coefficients in (42) and (43) have the following asymptotic property: $\hat{\beta}_n \rightarrow \beta$ and $\hat{\alpha}_n \rightarrow \alpha$ as $n \rightarrow \infty$.

By **lemma 5**, if there is a linear relationship between the noise variance and the spot volatility of a particular financial asset, the regression (43) can asymptotically discover it. Figure 1 and Figure 2 shows the least square regression plots for high-frequency transaction data in April, 2013 of 12 stocks in IT, financial, manufacturing and retailing industries.

Of course, one can investigate the statistical properties of this type of linear regression in more detail, and probably there are non-linear relations, these issues will be addressed in our future research.

7.2 Model extension: endogenous noise

As documented in Jacod et al. [2009], the identification assumption (6) is quite strong. The extension we discuss in the subsection is to abandon the identification assumption $Z_t = X_t, \forall t \in [0, T]$, as it turns out, the generalization of this type allows the market microstructure noise to be endogenous (noise is correlated with the efficient price).

Note that in model (1), (3) and (4), conditioning on latent variable $\omega^{(0)}$, ϵ_t is a mean-zero random variable, i.e., $\int_{\mathbb{R}} (y - Z_t) Q_t(X_t, dy) = 0$, put it in another way,

$$E(\epsilon_t | \omega^{(0)}) = E(Y_t - E(Y_t | \omega^{(0)})) | \omega^{(0)} = E(Y_t | \omega^{(0)}) - E(Y_t | \omega^{(0)}) = 0$$

however, the conditional mean of e_t is not necessarily 0:

$$E(e_t | \omega^{(0)}) = E(Y_t - X_t | \omega^{(0)}) = Z_t - X_t$$

This mechanism enables us to, non-parametrically, introduce endogenous noise into our model. We can allow instantaneous/realized correlation between the latent process $\{X_t\}_{t \geq 0}$ and the noise process $\{e_t\}_{t \geq 0}$. Take the instantaneous correlation $Cov(X_t, e_t)$ as an example, rather than assuming the conditional mean, instead assuming the unconditional mean of $Y_t - X_t$ is zero, i.e., $E_{\mathbb{P}}(Y_t - X_t) = 0$, observe that:

$$\begin{aligned} Cov(X_t, e_t) &= E_{\mathbb{P}} \{ (X_t - E_{\mathbb{P}} X_t) [(Y_t - X_t) - E_{\mathbb{P}}(Y_t - X_t)] \} \\ &= E_{\mathbb{P}} [X_t (Y_t - X_t)] \\ &= E_{\mathbb{P}^{(0)}} [X_t \cdot E_{\mathbb{P}^{(1)}}(Y_t - X_t | \omega^{(0)})] \end{aligned}$$

thus

$$\begin{aligned} Cov(X_t, e_t) &= \int_{\Omega^{(0)}} X_t(\omega^{(0)}) \left[\int_{\mathbb{R}} (y - X_t(\omega^{(0)})) Q_t(d\omega^{(0)}, dy) \right] \mathbb{P}^{(0)}(d\omega^{(0)}) \\ &= \int_{\Omega^{(0)}} \left[X_t(\omega^{(0)}) \int_{\mathbb{R}} y Q_t(\omega^{(0)}, dy) - X_t^2(\omega^{(0)}) \right] \mathbb{P}^{(0)}(d\omega^{(0)}) \\ &= \int_{\Omega^{(0)}} [X_t(\omega^{(0)}) (Z_t(\omega^{(0)}) - X_t(\omega^{(0)}))] \mathbb{P}^{(0)}(d\omega^{(0)}) \\ &= E_{\mathbb{P}^{(0)}} [X_t Z_t] - E_{\mathbb{P}^{(0)}} [X_t^2] \end{aligned}$$

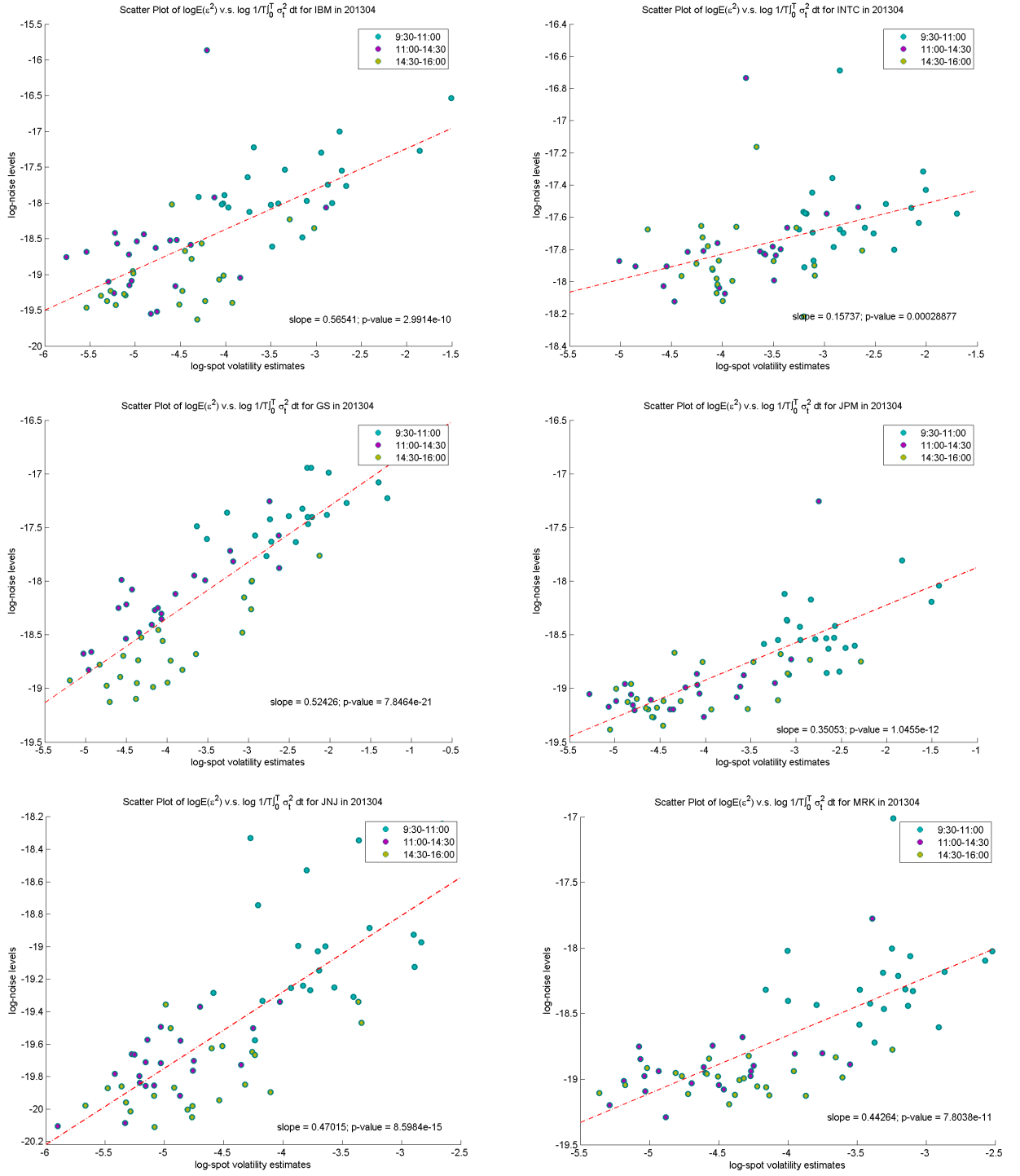


Figure 1: Scatter Plot of $\log \left(E(\widehat{\epsilon^2} | \omega^{(0)}) \right)$ against $\log \left(\frac{1}{t} \int_0^t \sigma_t^2 dt \right)$, t represents a particular period in each day (see the legends). The red dotted line is the fitted regression line. The upper panel exhibits the linear regression plots for IT companies, the middle panel exhibits the scatter plots for finance corporates, the lower panel exhibits the plot for medical and pharmaceutical companies.

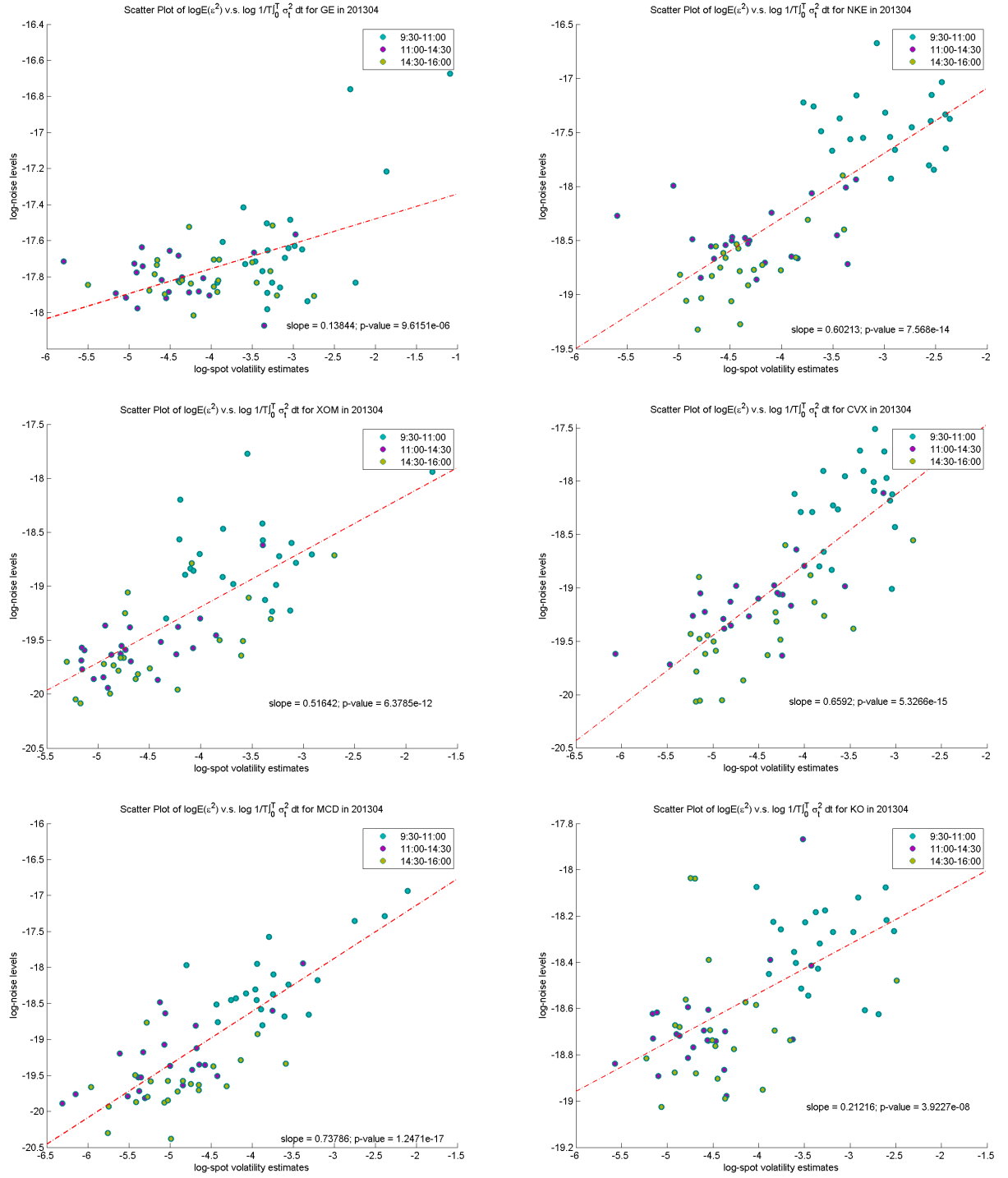


Figure 2: Scatter Plot of $\log \left(E(\epsilon^2 | \omega^{(0)}) \right)$ against $\log \left(\frac{1}{t} \int_0^t \sigma_t^2 dt \right)$, t represents a particular period in each day (see the legends). The red dotted line is the fitted regression line. The upper panel exhibits the linear regression plots for manufacturing companies, the middle panel exhibits the scatter plots for energy corporates, the lower panel exhibits the plot for fast food and retail companies.

In [Jacod et al., 2009], the authors assumed $Z_t = X_t$, so there is no correlation in their model. However, as long as $E_{\mathbb{P}^{(0)}}[X_t Z_t] \neq E_{\mathbb{P}^{(0)}}[X_t^2]$, there is correlation between the latent process $\{X\}_{t \geq 0}$ defined by (1) and the noise process $\{e_t\}_{t \geq 0}$ defined by (2).

Similarly, there is also a correlation structure between e_t and Z_t :

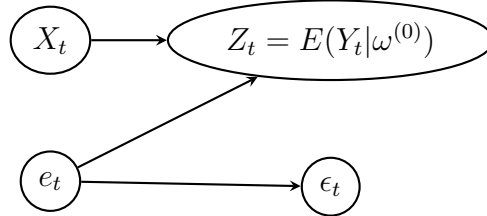
$$\text{Cov}(Z_t, e_t) = E_{\mathbb{P}^{(0)}}[Z_t^2] - E_{\mathbb{P}^{(0)}}[X_t Z_t] \quad (44)$$

However, by a similar fashion, it is not difficult to check that ϵ_t is not correlated with neither X_t nor Z_t ,

$$\text{Cov}(Z_t, \epsilon_t | \mathcal{F}_t) = \text{Cov}(X_t, \epsilon_t | \mathcal{F}_t) = 0$$

An intuitive interpretation is that e_t carries some relevant information about the processes defined on the latent probability space, so it is correlated with the latent random variables X_t and Z_t . In contrast, ϵ_t is a pure noise and conveys no useful information about the latent processes, the correlation between ϵ_t and any latent random variable is zero. In this example, we call $\{e_t\}_{t \geq 0}$ “*endogenous market microstructure noise*”.

In our model, we allow arbitrary fashion for the noise process up to the time-varying Markov kernel $Q_t(\cdot, \cdot)$. If the noise is correlated with the latent process, intuitively it is not hard to infer that $Z_t \neq X_t$. To put it in another way, $\{e_t\}_{t \geq 0}$ is not a pure noise process, the “noise” $\{e_t\}_{t \geq 0}$ conveys some information. The relationship can be viewed in the following picture, which shows both the latent variable X_t and endogenous noise e_t contribute to the estimable variable Z_t ; the remaining component of e_t which is not correlated with any latent variable is ϵ_t .



Remark From the assumptions of our model, when one tries to estimate the integrated volatility, the quantity which is actually estimated is $\langle Z, Z \rangle_T$, not necessarily the usually desired target $\langle X, X \rangle_T$. This point is discussed by [Li and Mykland, 2007]. In contrast to [Jacod et al., 2009], we do not assume $\int_{\mathbb{R}} y Q_t(\omega^{(0)}, dy) = X_t(\omega^{(0)})$. In other words, in the case where $Z_t \neq X_t$, the integrated volatility $\langle X, X \rangle_T$ of the latent process is not identifiable; however, if we are satisfied with estimating $\langle Z, Z \rangle_T$, then we are able to introduce some conditional correlation between the efficient price and market microstructure noise.

One conceptual finding from the model extension is the informational content in the market microstructure noise $\{e_t\}_{t \geq 0}$ with respect to the efficient price (or latent process in probabilistic term) modeled by Itô process $\{X_t\}_{t \geq 0}$.

The intuition behind this point comes from market microstructure theory [O'Hara, 1998, 2003]. As in the classical asset pricing theory, we take the price as given, and conduct trading and hedging strategies, portfolio allocation and risk management, while regarding the efficient prices as exogenous. But the price discovery and price formation depend on the behaviors of market participants, no price will be produced without investment activities of various market participants. It is the balance between demand and supply from investors, it is the psychology of people in the market, it is the synthesis of microscopic effects of behaviors of each participant in the market, that determine the prices. Thus the efficient price should be an endogenous process in the financial market. It is one of striking difference between asset pricing and market microstructure theory: the classical asset pricing theory assumes frictionless and competitive market in which people do not have to worry about the price impact and illiquidity. While, in market microstructure theory, the modelers need to look inside the “black box” of the trading processes, and take market making, price discovery, liquidity formation, inventory control, insider information into account.

Since we consider the price as endogenous, which, for example, affected by transaction costs (like bid-ask spread), inventory control, discrete adjustment of price, lagged incorporation of new information, insider trading and adverse selection brought by asymmetric information, lack of liquidity caused by one or several of the factors mentioned above, the Itô process is merely an approximation to the efficient price observed at high-frequency, at which market microstructure effects manifest itself to such extent that the accumulated noise swamps the integrated volatility of the latent Itô process and the variation in microstructure noise dominates the total variance.

Therefore, it is reasonable (even indispensable) to introduce an appropriate correlation structure into our model, at least from a point of view of microstructure theory, and for sake of modeling the prices in a low-latency and millisecond level. This topic is not the focus of this paper, in-depth discussion and treatment on endogenous market microstructure noise will be addressed in our future research.

8 Simulation

8.1 Simulation scenario

The data are generated from the Heston model [Heston, 1993]¹⁰:

$$dX_t = \sigma_t dW_t \quad (45)$$

$$d\sigma_t^2 = \kappa(\bar{\sigma}^2 - \sigma_t^2) + s\sigma_t dB_t \quad (46)$$

with $Cov(dW_t, dB_t) = \rho dt$.

Then we add stationary or non-stationary Gaussian noise to the latent semimartingale, the averaged variance of the noise is a^2 , which can be either stationary or non-stationary.

Table 1: The Parameters in Heston model for Monte Carlo simulation

X_0	κ	s	$\bar{\sigma}^2$	ρ	a
$\log(100)$	6	0.5	0.16	-0.6	0.05

The TSRV estimators are computed by taking $K = \lfloor n^{2/3} \rfloor$, and they have been adjusted for the finite-sample consideration:

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(TSRV, K), adj} &= \left(1 - \frac{n - K + 1}{nK}\right)^{-1} \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \\ \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K), adj} &= \left(1 - \frac{1}{K}\right)^{-1} \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} \end{aligned}$$

The non-stationary noise is either u-shaped or reversed u-shaped¹¹:

$$\epsilon_{t_i}^{(u)} \sim \left(\left(1 + 2\frac{i}{n}\right)^2 - 2\left(1 + 2\frac{i}{n}\right) + \frac{14}{3} \right) \times N(0, a^2) \quad (47)$$

$$\epsilon_{t_i}^{(ru)} \sim \left(-\left(1 + 2\frac{i}{n}\right)^2 + 2\left(1 + 2\frac{i}{n}\right) - \frac{11}{3} \right) \times N(0, a^2) \quad (48)$$

8.2 Simulation of the stationarity test 1: $N(Y, K, 2)_T^n$

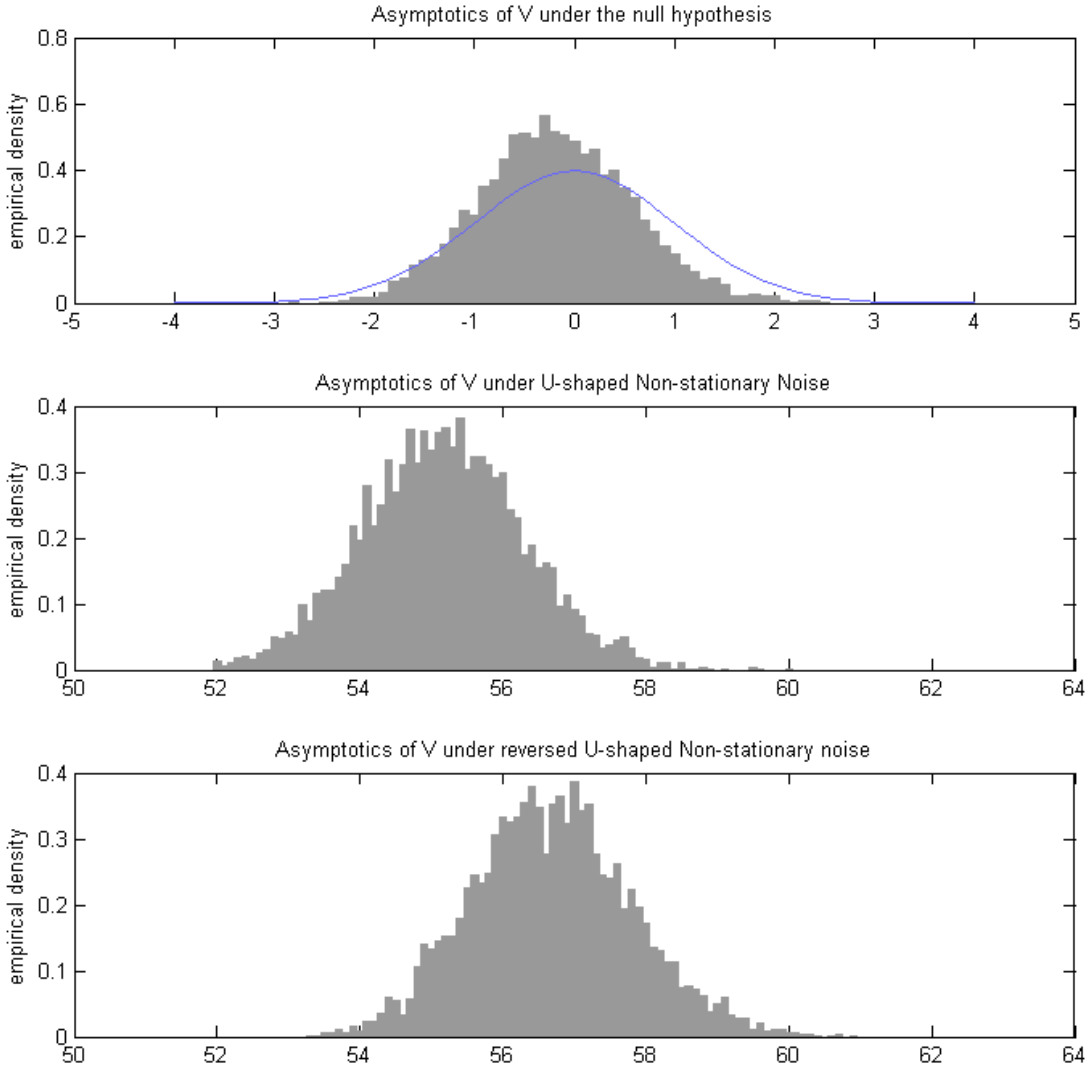
In Figure 3, we show the simulation results for the first test statistic $N(Y, K)_{[0, T]}$ on the time interval $[0, T]$, where we take T as one trading day. Our simulation is conducted

¹⁰The simulation scenario is adapted from Aït-Sahalia et al. [2010].

¹¹The functional forms of (47) and (48) are not meant to approximate the real intraday pattern of time variation in market microstructure noise variance. Instead, the symmetric U-shaped and reversed U-shaped curves are used to manifest the edge effects in TSRV and numerically collaborate theoretical results.

in three different circumstances: stationary noise, u-shaped noise (47), and reversed u-shaped noise (48). The plots shows the empirical distributions of the scaled $N(Y, K, 2)_T^n$ according to its asymptotic distribution (19), and we compare the empirical distribution with the density of $N(0, 1)$. The simulation results are consistent with the prediction given by our theory.

8.3 Simulation of the stationary test 2: $V(Y, K, 2)_T^n$ & $U(Y, K, 2)_T^n$



Asymptotic Distribution of $V(Y, K, 2)_T^n$ under stationary noise and U-shaped & reversed U-shaped heterogeneous noise. The blue line is the probability density function of standard normal $N(0, 1)$, the estimators computed from the simulated samples are shifted and scaled according to integrated volatilities, rate of convergence and the

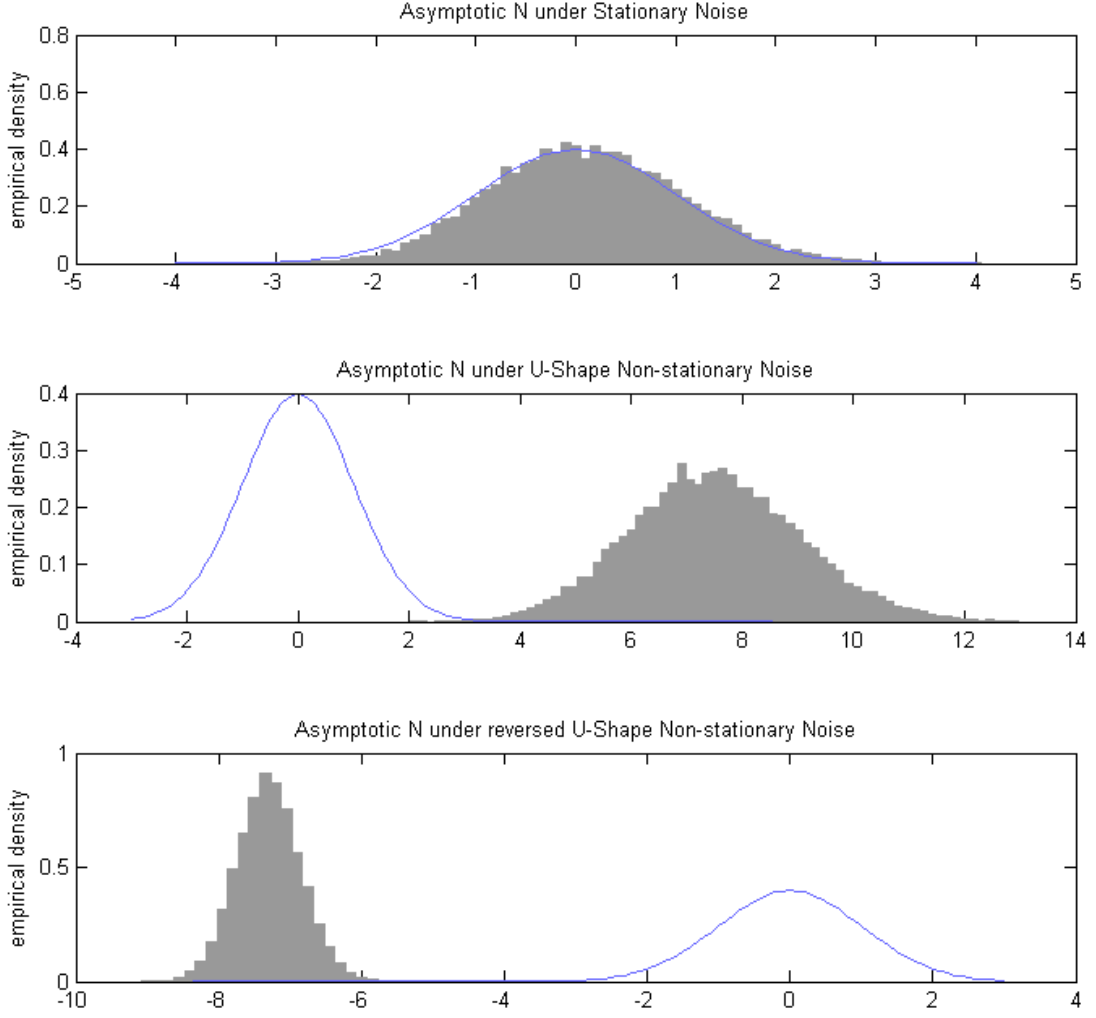
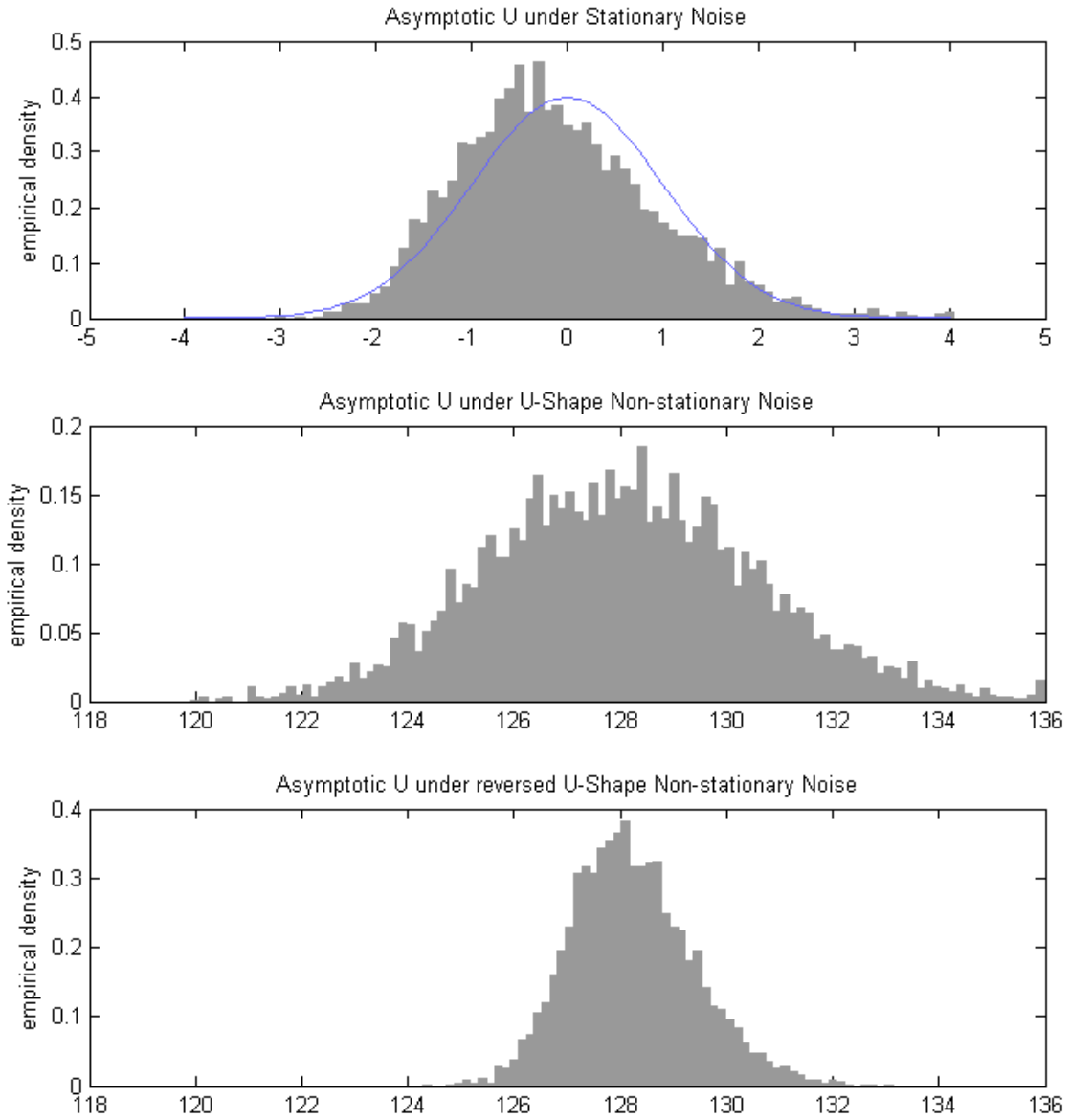


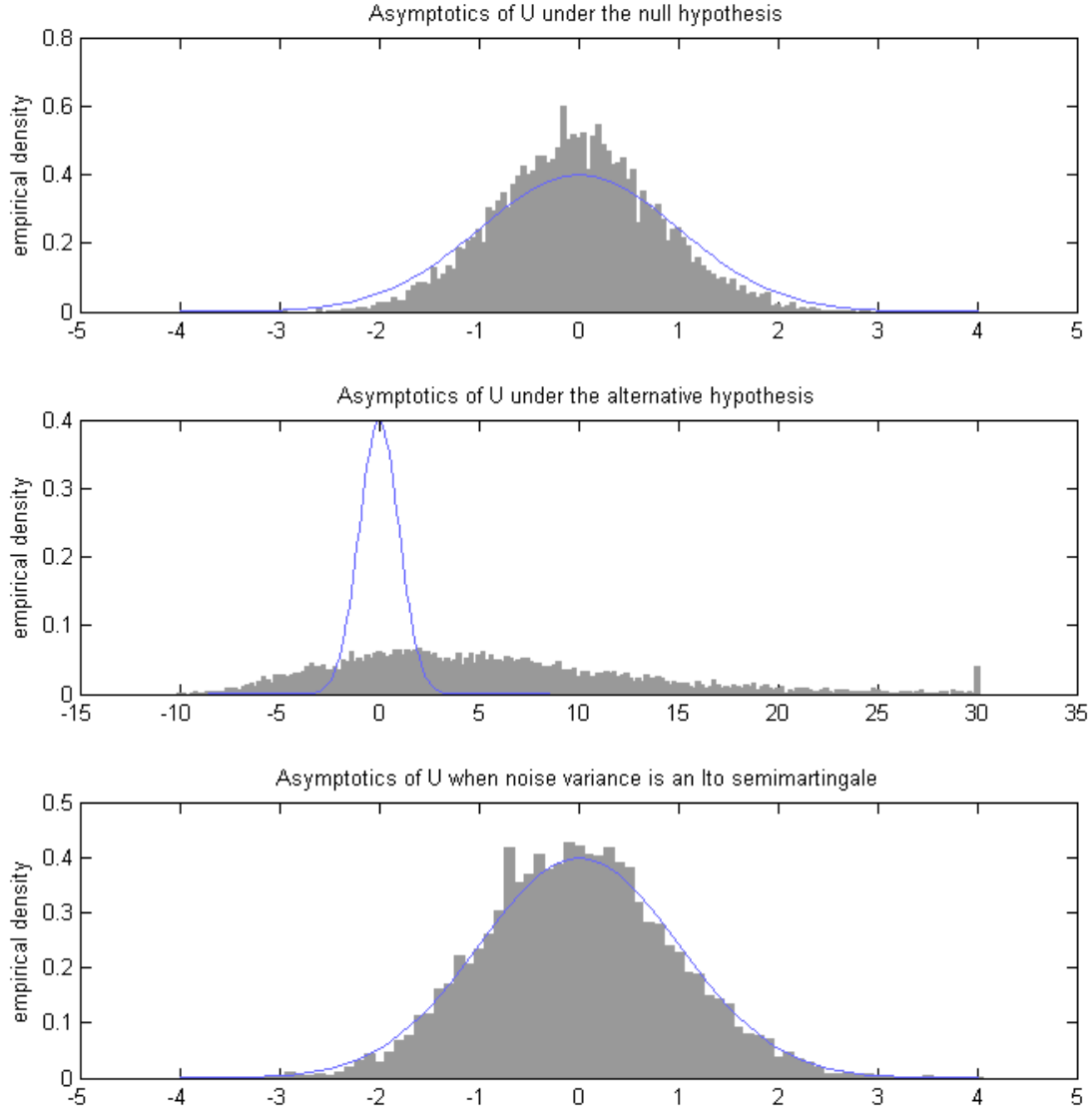
Figure 3: Asymptotic Distribution of $N(Y, K)_{[0,T]}$ under stationary noise and U-shaped & reversed U-shaped heterogeneous noise. The blue line is the probability density function of standard normal $N(0,1)$, the estimators computed from the simulated samples are shifted and scaled according to integrated volatilities, rate of convergence and the asymptotic variances.

asymptotic variances.



Asymptotic Distribution of $U(Y, K, 2)_T^n$ under stationary noise and U-shaped & reversed U-shaped heterogeneous noise. The blue line is the probability density function of standard normal $N(0, 1)$, the estimators computed from the simulated samples are shifted and scaled according to integrated volatilities, rate of convergence and the asymptotic variances.

8.4 Simulation of $U(Y, K, 2)_T^n$ when the noise variance is random



Asymptotic Distribution of $U(Y, K, 2)_T^n$ under stationary noise and heterogeneous noise (whose variance are modelled as an Itô process). The blue line is the probability density function of standard normal $N(0, 1)$, the estimators computed from the simulated samples are shifted and scaled according to the asymptotic distribution.

9 Empirical studies

9.1 Empirical evidence of non-stationary market microstructure noise

Using the realized variances computed from the fastest time scale in different time periods, we can obtain intra-day and daily estimates of the market microstructure noise level, then we can compare the noise levels of each stock across different time periods.

The upper panel in Figure 4 shows the intra-day pattern of market microstructure noises of different stocks, the lower panel exhibits daily variation in 2008. Figure 5 and Figure 6 show the intra-day variations in market microstructure noises of individual stocks in the first 4 months in 2013.

9.2 Empirical test results

In this subsection, we apply our tests onto the real high-frequency financial data from the TAQ data set in WRDS. We take several components in Dow Jones Industrial Average (DJIA30): Intel Corporation (INTC), International Business Machines Corporation (IBM), Goldman Sachs (GS), JPMorgan Chase (JPM), Exxon Mobil Corporation (XOM), and WalMart (WMT). We compute the test statistics and the p-values for these stocks during the 22 business days in April, 2013. Besides, in Figure 7 and Figure 8 we plot the whole trend of the test statistics $N(Y, K)_T^n$ and $V(Y, K, 2)_T^n$ during the period January 3, 2006 to December 31, 2013 as measures of liquidity.

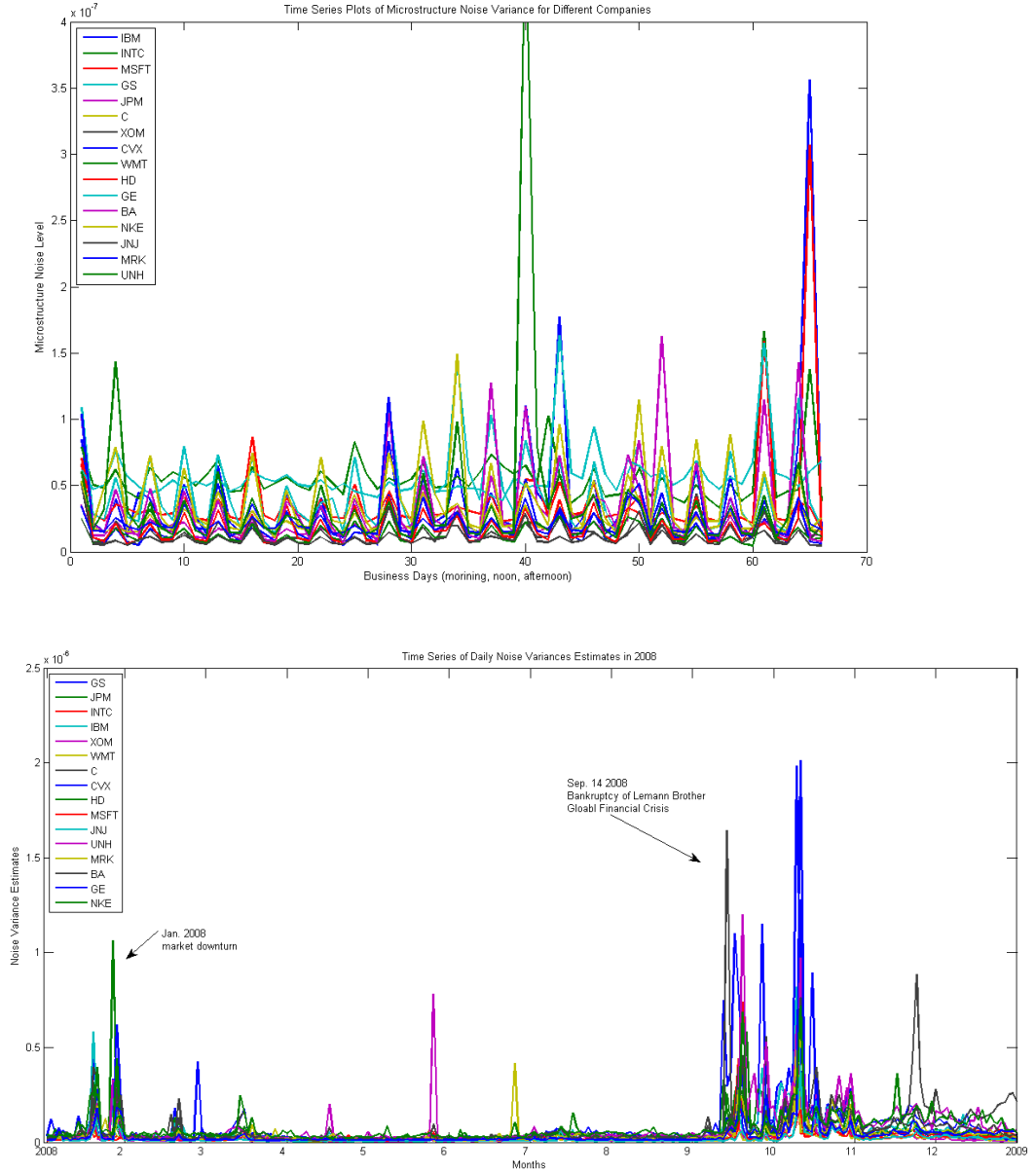


Figure 4: **Upper Panel:** intraday time Series of $E(\widehat{\epsilon^2}|\omega^{(0)}) = \frac{1}{2n}[Y, Y]_T^{(all)}$ for the three DJIA components in April, 2013. We divided each business day into three segments: 9:30-11:00, 11:00-14:30 and 14:30-16:00, and estimated $E(\epsilon^2|\omega^{(0)})$ in each segment for all the business days in April, 2013. There were 22 business days in that month, so we have 66 $E(\epsilon^2|\omega^{(0)})$ estimates for each company. The estimates of each company exhibits a diurnal pattern. **Lower Panel:** Daily noise variance estimates in 2008, with a simple event-history analysis. During the turmoil of financial crisis in 2008, the market microstructure noise surged up, the quality of the market worsened strikingly.

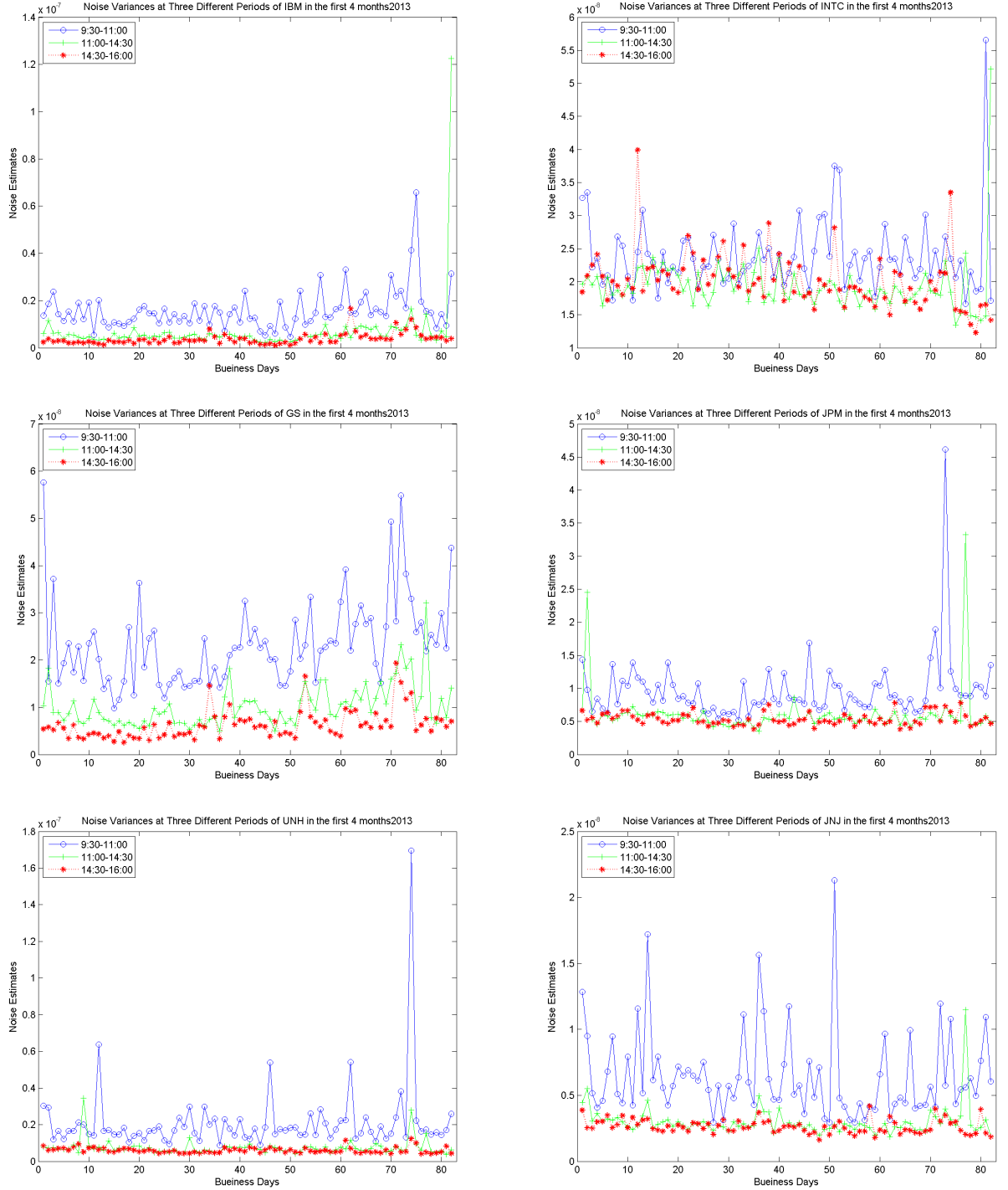


Figure 5: Time Series of $\widehat{E(\epsilon^2|\omega^{(0)})} = \frac{1}{2n}[Y, Y]_T^{(all)}$ in a particular period each day. For example, the **green line** is the time series plot of estimated noise level around noon (11:00-14:30) across different business days in April, 2013. The upper panel exhibits the microstructure noise time series of two IT companies: IBM and INTC, the middle panel exhibits the time series for two financial companies: GS and JPM, the lower panel exhibits the time series for medical and pharmaceutical companies: UNH and JNJ.

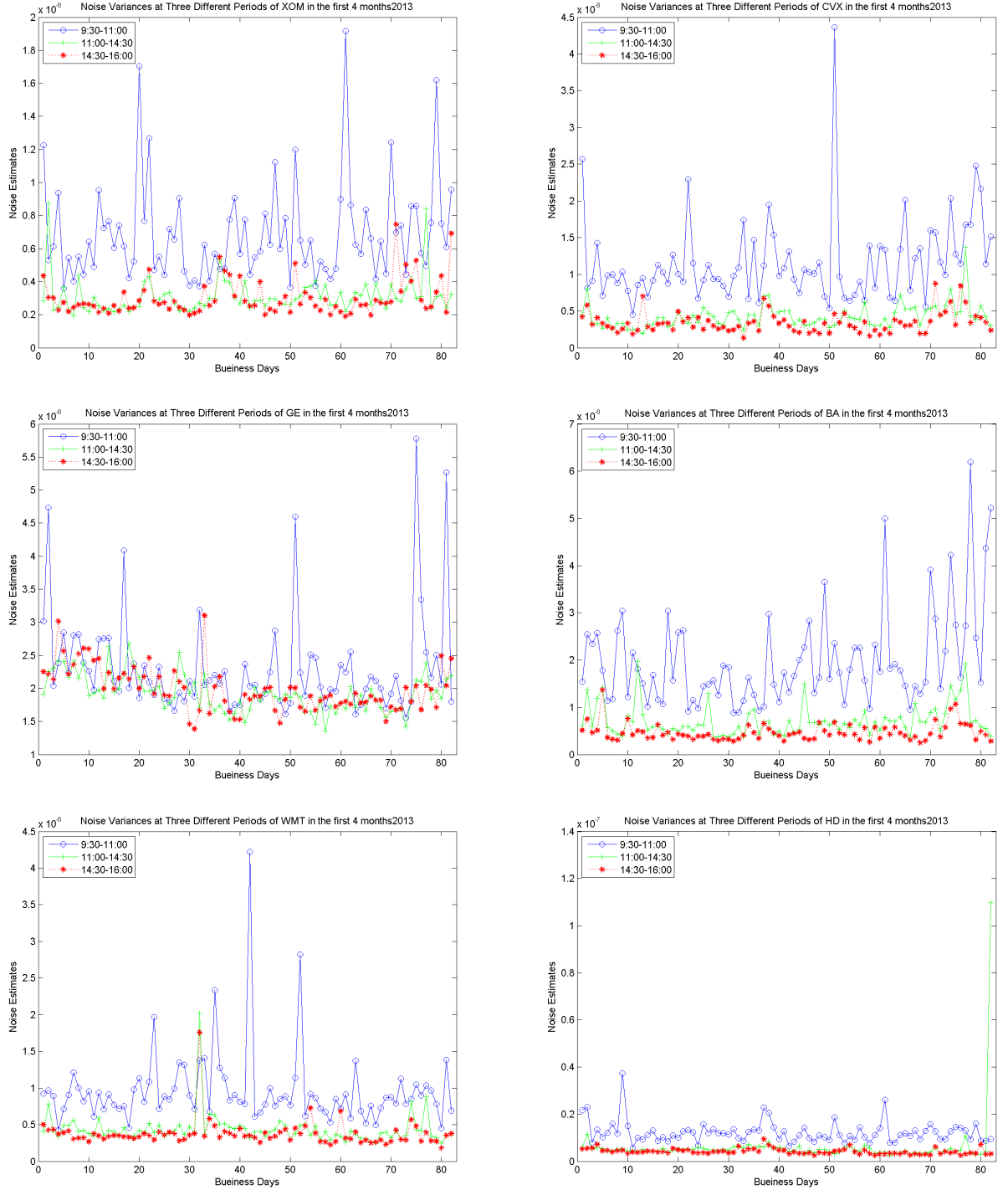


Figure 6: Time Series of $\widehat{E(\epsilon^2|\omega^{(0)})} = \frac{1}{2n}[Y, Y]_T^{(all)}$ in a particular period each day. For example, the green line is the time series plot of estimated noise level around noon (11:00-14:30) across different business days in April, 2013. The upper panel exhibits the time series of energy corporations (gasoline and oil): XOM and CVX, the middle panel exhibits the time series for manufacturing companies: GE and BA, the lower panel exhibits the time series of retailing companies: WMT and HD.

The Standardized First Test Statistics $N(Y, K)_T$ Computed from DJIA Components

Dates yyyy-mm-dd	DJIA Components									
	IBM		XOM		INTC		GS		GE	
	$N(Y, K)_T$	p-value	$N(Y, K)_T$	p-value	$N(Y, K)_T$	p-value	$N(Y, K)_T$	p-value	$N(Y, K)_T$	p-value
2013-04-01	0.5942	0.2762	6.0114	9.1947e-10	17.3676	0	0.9125	0.1807	6.4765	4.6925e-11
2013-04-02	3.8894	5.0246e-05	16.7202	0	12.3133	0	8.6813	0	26.2744	0
2013-04-03	6.8579	3.4941e-12	11.1238	0	12.6015	0	9.9089	0	4.5688	2.4529e-06
2013-04-04	4.5851	2.2690e-06	11.7737	0	11.8105	0	7.4771	3.7970e-14	8.3468	0
2013-04-05	8.6943	0	19.6103	0	21.9399	0	13.0797	0	7.7996	3.1086e-15
2013-04-08	12.0086	0	10.2720	0	19.6533	0	12.0044	0	8.7725	0
2013-04-09	4.4107	5.1507e-06	4.2196	1.0152e-05	14.5840	0	3.9217	4.3971e-05	2.8118	0.0025
2013-04-10	10.7967	0	20.3985	0	12.4934	0	1.4729	0.0704	12.1430	0
2013-04-11	10.5358	0	8.4332	0	19.8102	0	5.5467	1.4557e-08	7.6796	7.9936e-15
2013-04-12	9.8741	0	18.8744	0	9.7960	0	10.4689	0	11.3813	0
2013-04-15	8.6767	0	37.0635	0	11.8791	0	5.0028	2.8247e-07	6.6791	1.2023e-11
2013-04-16	11.5517	0	25.8213	0	11.0252	0	5.5612	1.3384e-08	16.5744	0
2013-04-17	11.2338	0	4.2163	1.2419e-05	20.6048	0	5.5168	1.7261e-08	13.6559	0
2013-04-18	15.1748	0	14.7396	0	49.2313	0	3.1477	8.2284e-04	10.6347	0
2013-04-19	29.7852	0	18.3013	0	10.8806	0	9.7611	0	18.5074	0
2013-04-22	13.4899	0	7.0150	1.1479e-12	10.0430	0	7.5659	1.9207e-14	12.9960	0
2013-04-23	11.0911	0	0.9798	0.1636	1.5144	0.065	0.4083	0.3415	26.3066	0
2013-04-24	10.6420	0	26.4967	0	22.6824	0	9.6762	0	20.8122	0
2013-04-25	12.8092	0	13.4558	0	15.4190	0	9.4322	0	9.3956	0
2013-04-26	7.1480	4.4031e-13	14.8469	0	15.1904	0	3.0896	0.0010	6.0681	6.4723e-10
2013-04-29	3.4021	3.3438e-04	19.2697	0	0.8441	0.1993	9.3275	0	-0.0481	0.4808
2013-04-30	0.4047	0.3428	12.8344	0	-0.2676	0.3945	10.1050	0	7.4785	3.7637e-14

The Standardized Second Test Statistics $V(Y, K, 2)_T$ Computed from DJIA Components

Dates yyyy-mm-dd	DJIA Components									
	IBM		XOM		INTC		GS		GE	
	V	p-value	V	p-value	V	p-value	V	p-value	V	p-value
2013-04-01	10.9938	0	27.0149	0	21.4302	0	10.1113	0	30.4807	0
2013-04-02	3.2344	6.0958e-04	24.8188	0	9.3638	0	7.6711	8.5487e-15	38.4981	0
2013-04-03	1.9156	0.0277	12.0170	0	18.3532	0	4.3272	7.5517e-06	7.5493	2.1871e-14
2013-04-04	5.8728	2.1431e-09	13.3469	0	14.6418	0	6.0005	9.8349e-10	14.0090	0
2013-04-05	7.6103	1.3656e-14	20.0585	0	34.8890	0	11.4651	0	3.7583	8.5532e-05
2013-04-08	13.3551	0	4.5174	3.1307e-06	30.4338	0	9.0230	0	13.9205	0
2013-04-09	1.4434	0.0745	3.3262	4.4016e-04	16.7477	0	2.8457	0.0022	8.3337	0
2013-04-10	4.4848	3.6493e-06	25.0742	0	8.6673	0	3.2815	5.1624e-04	10.4242	0
2013-04-11	4.3129	8.0553e-06	4.6223	1.8976e-06	34.6739	0	4.0614	2.4387e-05	3.4916	2.4005e-04
2013-04-12	9.0331	0	33.4960	0	7.5215	2.7089e-14	4.6699	1.5067e-06	16.9798	0
2013-04-15	5.7864	3.5948e-09	37.6814	0	7.7029	6.6613e-15	21.0097	0	21.4712	0
2013-04-16	9.3722	0	26.5122	0	14.9960	0	4.9056	4.6576e-07	16.3368	0
2013-04-17	11.6865	0	10.0344	0	30.9887	0	2.1944	0.0141	11.6983	0
2013-04-18	13.2016	0	14.6304	0	135.0731	0	1.1928	0.1165	8.8386	0
2013-04-19	64.5239	0	44.9676	0	10.1527	0	7.5036	3.1086e-14	17.1892	0
2013-04-22	16.8527	0	5.4702	2.2472e-08	27.0663	0	7.3061	1.3756e-13	80.9389	0
2013-04-23	44.0966	0	45.6640	0	9.3016	0	41.7062	0	62.0387	0
2013-04-24	3.7373	9.3002e-05	44.4829	0	41.2661	0	4.2316	1.1602e-05	26.8806	0
2013-04-25	4.9287	4.1380e-07	18.6223	0	19.3681	0	6.2980	1.5079e-10	10.9943	0
2013-04-26	1.7234	0.0424	12.8827	0	22.0351	0	0.1381	0.4451	16.7115	0
2013-04-29	2.3578	0.0092	25.3177	0	56.7234	0	4.6938	1.3412e-06	17.4994	0
2013-04-30	40.0189	0	11.3305	0	24.0035	0	7.0859	6.9078e-13	16.9568	0

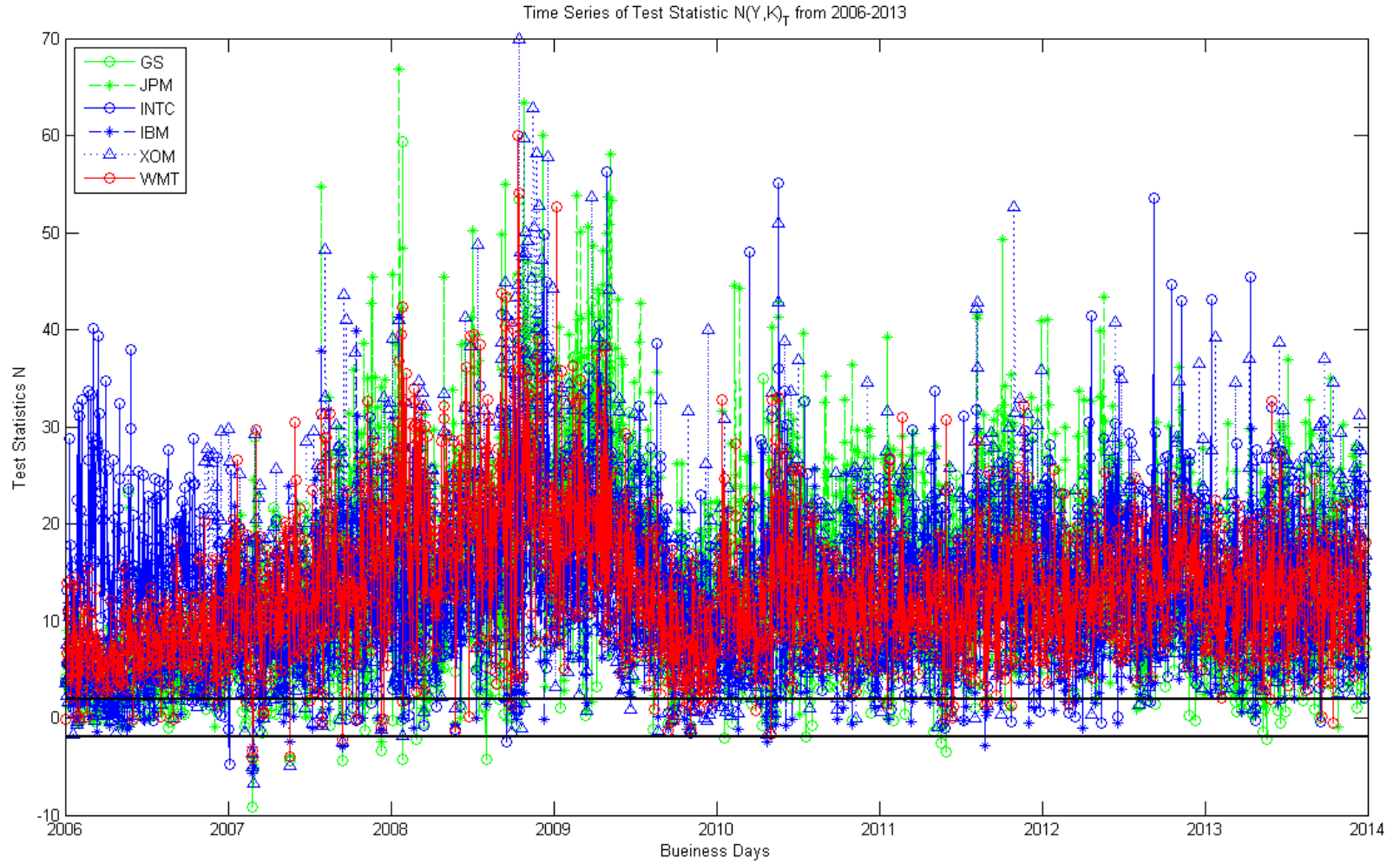


Figure 7: The six time series of the standardized first test statistic $N(Y, K)_T^n \equiv \frac{\sqrt{K} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - \widehat{\langle X, X \rangle}_T^{(TSRV, K)} \right)}{\sqrt{2E(\epsilon^4)}}$ computed daily using intra-day ultrahigh frequency financial data. The black horizontal lines around zero are .025% and .975% quantiles of the standard normal distribution.

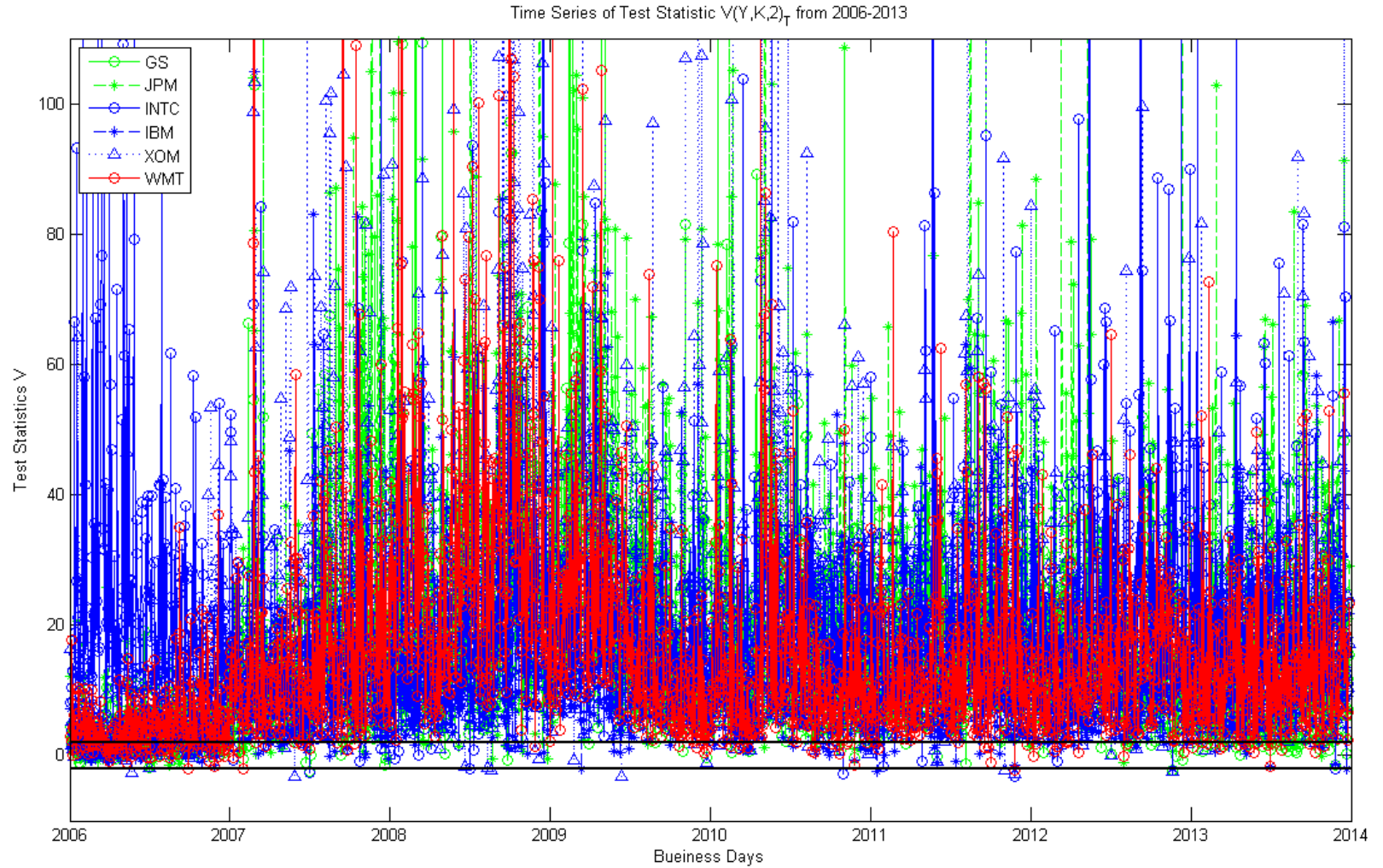


Figure 8: The six time series of the standardized second test statistic $V(Y,K,2)_T^n \equiv \frac{\sqrt{r_n - s_n + 1} \left(V(Y,K,2)_T^n - \frac{1}{n} \left([Y,Y,Y,Y]_T^{(all)} - \frac{3}{2n} ([Y,Y]_T^{(all)})^2 \right) \right)}{\sqrt{\hat{\eta}^2}}$ computed daily using intra-day ultrahigh frequency financial data. The black horizontal lines around zero are .025% and .975% quantile of the standard normal distribution.

10 Conclusion

In this paper, we use the model with a general form for microstructure noise, allowing the variance of the noise $\{\epsilon_t\}_{t \geq 0}$ to depend on latent variables, such as $\{X_t(\omega^{(0)})\}_{t \geq 0}$ and $\{\sigma_t^2(\omega^{(0)})\}_{t \geq 0}$ and the time t . In particular, we discuss the implication of the non-stationarity in noise for integrated volatility estimation, in which the market microstructure noise level is time-varying and the time-dependence may arise from the direct relation with the time, or from the indirect relation via the latent process itself as well as the latent volatility of the latent process.

We studied the behavior of TSRV under the contamination of non-stationary noise. Like [Kalnina and Linton \[2008\]](#), we find that the TSRV suffers from an bias of the same magnitude as its asymptotic variance. The noise level in the morning is higher than those at the noon and in the afternoon in a typical business day. To overcome the difficulty brought by the non-stationary market microstructure noise, we use a modified version of TSRV [[Kalnina and Linton, 2008](#)] by which we could eliminate the edge effect due to any non-stationary noise.

Based on the remedy for non-stationary market microstructure noise, we can exploit the edge effect and test the stationarity/non-stationarity in high-frequency data based on asymptotic behaviors: some functionals of nonparametric estimator of volatility and noise variance obey stable central limit theorems under the null hypothesis that the noise is stationary, and our test statistics explode in the in-fill asymptotics when the noise is non-stationary. In particular, we designed test statistics, one is $N(Y, K)_T^n$, the other is $V(Y, K, 2)_T^n$. $N(Y, K)_T^n$ takes advantage the edge effect by comparing the noise at the edges and that in the middle; $V(Y, K, 2)_T^n$ is built upon the first test statistics computed in different time windows. The first test statistic is computationally convenient, and the second test enjoys better statistical property in term of approximation accuracy, but it is more computationally expensive. We suggest one can choose between these two tests based on the specific need.

Furthermore, we find an asymptotically equivalent statistic $U(Y, K, 2)_T^n$ of $V(Y, K, 2)_T^n$ when the market microstructure noise is stationary. After being scaled by a constant, $U(Y, K, 2)_T^n$ can consistently estimate the quadratic variation of the variance process of the market microstructure noise when the noise is non-stationary. So, besides testing whether the market microstructure noise is stationary, $U(Y, K, 2)_T^n$ can also measure aggregate liquidity risk using high frequency data.

To verify the relevance of our general model, we analyze the real high-frequency financial data from NYSE. As the data analysis of DJIA components from 2006-2013 showed, the variances of market microstructure noise have indeed changed recently, both daily and intra-daily, which agrees with the empirical results in the literature. Besides, we find that the timing of the sudden increase in noise variance in Sep. 2008 coincided with the beginning of the global financial crisis triggered by mortgage sub-prime crisis started from 2007. Furthermore, market microstructure noise could be a good measure of the market quality (market liquidity, market depth etc.) [[Hasbrouck, 1993](#), [O'Hara, 2003](#), [Aït-Sahalia and Yu, 2009](#)], using our test statistic, we can measure

the liquidity risk in individual stocks and can also evaluate the overall liquidity risk of the financial markets. The time series of our test statistics show that our test statistics can disclose a pattern during the financial crisis 2008-2009 which indicating an increase in the intra-day transaction costs in the financial markets.

References

- Admati, A. R. and Pfleiderer, P. (1988). A theory of intraday patterns: volume and price variability. *The review of financial studies*, 1(1):3–40.
- Aït-Sahalia, Y., Fan, J., and Xiu, D. (2010). High-frequency covariance estimates with noisy and asynchronous financial data. *Journal of American Statistical Association*, 105(492):1504–1517.
- Aït-Sahalia, Y. and Jacod, J. (2014). High-Frequency financial econometrics. *Princeton University Press*.
- Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *The Review of Financial Studies*, 18(2):351–416.
- Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2006). Comment on “Realized variance and market microstructure noise”. *Journal of Business & Economic Statistics*, 24(2):162–167.
- Aït-Sahalia, Y., Mykland, P. A., and Zhang, L. (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics*, 160.
- Aït-Sahalia, Y. and Yu, J. (2009). High Frequency market microstructure noise estimates and liquidity measures. *The Annals of Applied Statistics*, 3(1):422–457.
- Andersen, T. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, 96(453):42–55.
- Awartani, B., Corradi, V., and Distaso, W. (2009). Assessing market microstructure effects via realized volatility measures with an application to the Dow Jones Industrial Average stocks. *Journal of Business & Economic Statistics*, 27(2):251–265.
- Bandi, F., Russell, J., and Yang, C. (2013). Realized volatility forecasting in the presence of time-varying noise. *Journal of Business & Economic Statistics*, 31(3):331–345.
- Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility. *Journal of Financial Economics*.
- Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76(6):1481–1536.

- Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society*, B(64):253–280.
- Bibinger, M., Hautsch, N., Malec, P., and Reiß, M. (2014). Estimating the quadratic covariation matrix from noisy observations: local method of moments and efficiency. *The Annals of Statistics*, 42(4):1312–1346.
- Delbaen, F. and Schachermayer, W. (1994). A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*.
- Engle, R. F. (2000). The Econometrics of Ultra-High-Frequency Data. *Econometrica*, 68(1):1–22.
- Foster, D. and Nelson, D. (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica*, 64(64):139–174.
- Glosten, L. and Milgrom, P. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, 14.
- Gouriéroux, C., Jasiak, J., and Fol, G. L. (1999). Intra-day market activity. *Journal of Financial Markets*, (2):193–226.
- Harrison, M. and Pliska, S. (1981). Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Applications*.
- Hasbrouck, J. (1993). Assessing the quality of a security market: a new approach to transaction-cost measurement. *The Review of Financial Studies*, 6(1):191–212.
- Hasbrouck, J. (2007). Empirical market microstructure: the institutions, economics, and econometrics of securities trading. *Oxford University Press*.
- Hautsch, N. (2012). Econometrics of Financial High-Frequency Data. *Springer-Verlag Berlin Heidelberg*.
- Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343.
- Jacod, J. (1994). Limit of random measures associated with the increments of a Brownian semimartingale. *Tech. Rep. Université de Paris II*.
- Jacod, J. (2012). Statistics and High Frequency Data. *Statistical Methods for Stochastic Differential Equations, Chapter 3*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Book 124):191–310.
- Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and their Application*, 119.

- Jacod, J., Podolskij, M., and Vetter, M. (2010). Limit theorems for moving averages of discretized processes plus noise. *The Annals of Statistics*, 38(3):1478–1545.
- Jacod, J. and Protter, P. (1998). Asymptotic error distributions for the Euler method for stochastic differential equations. *Annals of Probability*, 26(26):267–3–7.
- Jacod, J. and Protter, P. (2012). Discretization of Processes. *Springer-Verlag*, Stochastic Modelling and Applied Probability (67).
- Jacod, J. and Shiryaev, A. (2003). Limit theorems for stochastic processes, 2nd edition. *Springer-Verlag*, Grundlehren der mathematischen Wissenschaften (288).
- Kalnina, I. and Linton, O. (2008). Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of Econometrics*, 147.
- Kyle, A. (1985). Continuous auctions and insider trading. *Econometrica*, 53(6):1315–1335.
- Li, Y. and Mykland, P. A. (2007). Are volatility estimators robust with respect to modeling assumption? *Bernoulli*, 13(3):601–622.
- Mykland, P. A. and Zhang, L. (2006). ANOVA for diffusions and Itô processes. *The Annals of Statistics*, 34(4):1931–1963.
- Mykland, P. A. and Zhang, L. (2012). The econometrics of high frequency data. *Statistical Methods for Stochastic Differential Equations*, Chapman & Hall/CRC Monographs on Statistics & Applied Probability (Book 124), Chapman and Hall/CRC(Chapter 2):109–190.
- Mykland, P. A. and Zhang, L. (2015a). Assessment of uncertainty in high frequency data: the observed asymptotic variance. *Working paper*.
- Mykland, P. A. and Zhang, L. (2015b). Between data cleaning and inference: pre-averaging and robust estimators of the efficient price. *Working paper*.
- O’Hara, M. (1998). Market microstructure theory. *Wiley*.
- O’Hara, M. (2003). Liquidity and price discovery. *The Journal of Finance*, LVIII(4):1335–1354.
- Podolskij, M. and Vetter, M. (2009). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3):634–658.
- Russell, J. R. and Engle, R. F. (2010). Analysis of High-Frequency Data. *Handbook of Financial Econometrics. Volume 1 - Tools and Techniques*, Handbooks in Finance, North-Holland:383–426.
- Stoll, H. R. (2000). Friction. *The Journal of Finance*, LV(4):1479–1514.

- Xiu, D. (2010). Quasi-maximum likelihood estimation of volatility with high frequency data. *Journal of Econometrics*, 159.
- Zhang, L. (2001). From martingales to ANOVA: implied and realized volatility. *Ph.D. Thesis, The University of Chicago*.
- Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli*, 12(6):1019–1043.
- Zhang, L., Mykland, P. A., and Aït-Sahalia, Y. (2005). A tale two time scales: determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association*, 100(472):1394–1411.

11 Appendix

11.1 Derivation of the edge effects of TSRV

All the calculations are conditional on the whole latent process X . Assuming the **Proposition 1** and **Lemma 6** in [Li and Mykland, 2007]:

Proposition 1. Assume that $E(|A_n||\omega^{(0)})$ is $O_p(1)$. Then A_n is $O_p(1)$.

Lemma 6. $M_T^{(2)}$ and $M_T^{(3)}$ are asymptotically independently normal conditional on the latent variable(s), both with variance $\frac{1}{T} \int_0^T (g_t(\omega^{(0)}))^2 dt$.

Define:

$$\begin{aligned} R_1 &= (\epsilon_{t_0}^2 - E(\epsilon_{t_0}^2 | \omega^{(0)})) + (\epsilon_{t_n}^2 - E(\epsilon_{t_n}^2 | \omega^{(0)})) \\ R_2 &= \sum_{k=1}^K \left[(\epsilon_{\min \mathcal{G}^{(k)}}^2 - E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)})) + (\epsilon_{\max \mathcal{G}^{(k)}}^2 - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \right] \end{aligned}$$

By assumptions of our model in section 2, we know:

$$E(R_1^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)}) = O_p(1)$$

where $\tau_l = \inf\{t : |X_t| \wedge \sigma_t > l\}$.

Thus, by **Proposition 1** and the fact that $P(\tau_l < T) \rightarrow 0$ as $l \rightarrow \infty$, we know $R_1 = O_p(1)$. Similarly,

$$\begin{aligned} &E(R_2^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)}) \\ &= E \left(\left(\sum_{k=1}^K \left[(\epsilon_{\min \mathcal{G}^{(k)}}^2 - E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)})) + (\epsilon_{\max \mathcal{G}^{(k)}}^2 - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \right] \right)^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)} \right) \\ &= \sum_{k=1}^K E \left(\left[(\epsilon_{\min \mathcal{G}^{(k)}}^2 - E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)})) + (\epsilon_{\max \mathcal{G}^{(k)}}^2 - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \right]^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)} \right) \\ &= O_p(K) \end{aligned}$$

the second equality holds since we assume Y_{t_0}, \dots, Y_{t_n} are conditionally independent given the X process, so as the noise terms defined by $\epsilon_{t_i} = Y_{t_i} - Z_{t_i}$, hence, $R_2 = O_p(\sqrt{K})$.

As a consequence,

$$\begin{aligned} [\epsilon, \epsilon]_T^{(all)} &= \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 = \sum_{i=1}^n (\epsilon_{t_i}^2 + \epsilon_{t_{i-1}}^2 - 2\epsilon_{t_i} \epsilon_{t_{i-1}}) \\ &= 2 \sum_{i=0}^n (\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)})) - 2 \sum_{i=1}^n \epsilon_{t_i} \epsilon_{t_{i-1}} + 2 \sum_{i=0}^n E(\epsilon_{t_i}^2 | \omega^{(0)}) - \epsilon_{t_0}^2 - \epsilon_{t_n}^2 \end{aligned}$$

thus

$$\begin{aligned} [\epsilon, \epsilon]_T^{(all)} &= \underbrace{2 \sum_{i=0}^n (\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)})) - 2 \sum_{i=1}^n \epsilon_{t_i} \epsilon_{t_{i-1}} + 2 \sum_{i=0}^n E(\epsilon_{t_i}^2 | \omega^{(0)})}_{2\sqrt{n}(M_T^{(1)} - M_T^{(2)})} \\ &\quad - \underbrace{\left(E(\epsilon_{t_0}^2 | \omega^{(0)}) + E(\epsilon_{t_n}^2 | \omega^{(0)}) \right)}_{O_p(1)} - \left[\underbrace{\left(\epsilon_{t_0}^2 - E(\epsilon_{t_0}^2 | \omega^{(0)}) \right) + \left(\epsilon_{t_n}^2 - E(\epsilon_{t_n}^2 | \omega^{(0)}) \right)}_{R_1 = O_p(1)} \right] \\ &= 2\sqrt{n} (M_T^{(1)} - M_T^{(2)}) + 2 \sum_{i=0}^n E(\epsilon_{t_i}^2 | \omega^{(0)}) + O_p(1) \end{aligned} \tag{49}$$

Furthermore,

$$\begin{aligned}
K[\epsilon, \epsilon]_T^{(avg, K)} &= \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} (\epsilon_{t_i} - \epsilon_{t_i, -})^2 = \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} (\epsilon_{t_i}^2 + \epsilon_{t_i, -}^2 - 2\epsilon_{t_i} \epsilon_{t_i, -}) \\
&= -2 \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_i, -} + \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i}^2 + \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i, -}^2 \\
&= 2 \left[\sum_{i=0}^n \epsilon_{t_i}^2 - \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_i, -} \right] - \sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2) \\
&= 2 \underbrace{\left[\sum_{i=0}^n (\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)})) - \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_i, -} \right]}_{\sqrt{n}(M^{(1)} - M^{(3)})} \\
&\quad + 2 \sum_{i=0}^n E(\epsilon_{t_i}^2 | \omega^{(0)}) - \sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2)
\end{aligned}$$

thus

$$\begin{aligned}
K[\epsilon, \epsilon]_T^{(avg, K)} &= 2\sqrt{n} (M_T^{(1)} - M_T^{(3)}) + 2 \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) - \sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2) \\
&\quad + 2 \sum_{k=1}^K E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \\
&= 2\sqrt{n} (M_T^{(1)} - M_T^{(3)}) + 2 \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) - \sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2) \\
&\quad + \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) + \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}))
\end{aligned}$$

Note in the RHS of the last equality above¹²,

$$\begin{aligned}
& - \sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2) + \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \\
&= - \left(\sum_{k=1}^K (\epsilon_{\min \mathcal{G}^{(k)}}^2 + \epsilon_{\max \mathcal{G}^{(k)}}^2) - \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \right) \\
&= - \sum_{k=1}^K [(\epsilon_{\min \mathcal{G}^{(k)}}^2 - E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)})) + (\epsilon_{\max \mathcal{G}^{(k)}}^2 - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}))] = -R_2
\end{aligned}$$

So, we get

$$\begin{aligned}
K[\epsilon, \epsilon]_T^{(avg, K)} &= 2\sqrt{n} (M_T^{(1)} - M_T^{(3)}) + 2 \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) - R_2 \\
&\quad + \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}))
\end{aligned} \tag{50}$$

¹²One thing must be noticed is that what does the “ $t_i \in \mathcal{G}^{(k)}$ ” really mean. Since the summation is over the summand $\epsilon_{t_i} - \epsilon_{t_i, -}$ for $t_i \in \mathcal{G}^{(k)}$, so the summation denote by $t_i \in \mathcal{G}^{(k)}$ is over the points set $\{\min \mathcal{G}^{(k)} + 1, \min \mathcal{G}^{(k)} + 1, \dots, \max \mathcal{G}^{(k)}\}$. The same thing applies to $t_i \in \mathcal{G}$ in the definition of $M_T^{(2)}$.

Thus, we have:

$$[\epsilon, \epsilon]_T^{(all)} - 2\sqrt{n} \left(M_T^{(1)} - M_T^{(2)} \right) = 2 \sum_{i=0}^n E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + O_p(1) \quad (51)$$

$$\begin{aligned} K[\epsilon, \epsilon]_T^{(avg, K)} - 2\sqrt{n} \left(M_T^{(1)} - M_T^{(3)} \right) &= 2 \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + O_p(\sqrt{K}) \\ &+ \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) \end{aligned} \quad (52)$$

the last equality holds because $R_2 = O_p(\sqrt{K})$.

By **Lemma 1**,

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(TSRV, K)} &= [Y, Y]_T^{(avg, K)} - \frac{\bar{n}}{n} [Y, Y]_T^{(all)} \\ &= \left([Z, Z]_T^{(avg, K)} + [\epsilon, \epsilon]_T^{(avg, K)} + O_p \left(\frac{1}{\sqrt{K}} \right) \right) - \underbrace{\left(\frac{\bar{n}}{n} [\epsilon, \epsilon]_T^{(all)} + O_p \left(\frac{\bar{n}}{n} \right) \right)}_{O_p \left(\frac{1}{K} \right)} \\ &= [Z, Z]_T^{(avg, K)} + [\epsilon, \epsilon]_T^{(avg, K)} - \underbrace{\frac{\bar{n}}{n} [\epsilon, \epsilon]_T^{(all)}}_{O_p \left(\frac{\sqrt{n}}{K} \right)} + O_p \left(\frac{1}{\sqrt{K}} \right) \end{aligned}$$

So, the convergence rate of $\widehat{\langle X, X \rangle}_T^{(TSRV, K)}$ to $[Z, Z]_T^{(avg, K)}$ is $\frac{K}{\sqrt{n}}$. So the TSRV estimator is consistent in this situation, next step is to find the limiting law, which requires a deeper investigation to the error term $[\epsilon, \epsilon]_T^{(avg, K)} - \frac{\bar{n}}{n} [\epsilon, \epsilon]_T^{(all)}$.

$$\begin{aligned} \widehat{\langle X, X \rangle}_T^{(TSRV, K)} - [Z, Z]_T^{(avg, K)} &= [\epsilon, \epsilon]_T^{(avg, K)} - \frac{\bar{n}}{n} [\epsilon, \epsilon]_T^{(all)} + O_p \left(\frac{1}{\sqrt{K}} \right) \\ &= \frac{2\sqrt{n}}{K} \left(M_T^{(1)} - M_T^{(3)} \right) + \frac{2}{K} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) - \frac{\bar{n}}{n} \left[2\sqrt{n} \left(M_T^{(1)} - M_T^{(2)} \right) + 2 \sum_{i=0}^n E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) \right] \\ &+ \frac{1}{K} \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \\ &= \frac{2\sqrt{n}}{K} \left(M_T^{(1)} - M_T^{(3)} \right) - 2 \cdot \frac{n-K+1}{K\sqrt{n}} \left(M_T^{(1)} - M_T^{(2)} \right) \\ &+ 2 \left[\frac{1}{K} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) - \frac{n-K+1}{nK} \sum_{i=0}^n E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) \right] \\ &+ \frac{1}{K} \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \end{aligned}$$

Observe that:

$$\begin{aligned} \frac{\sqrt{n}}{K} \left(M_T^{(1)} - M_T^{(3)} \right) - \frac{n-K+1}{K\sqrt{n}} \left(M_T^{(1)} - M_T^{(2)} \right) \\ = \underbrace{\frac{K-1}{\sqrt{n}K}}_{O_p \left(\frac{1}{\sqrt{n}} \right)} \cdot M_T^{(1)} + \underbrace{\frac{n-K+1}{\sqrt{n}K}}_{O_p \left(\frac{\sqrt{n}}{K} \right)} \cdot M_T^{(2)} - \frac{\sqrt{n}}{K} \cdot M_T^{(3)} \end{aligned}$$

And

$$\begin{aligned}
& \frac{2}{K} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) - 2 \cdot \frac{n-K+1}{nK} \sum_{i=0}^n E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + \frac{1}{K} \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) \\
&= \frac{2}{K} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) - 2 \cdot \frac{n-K+1}{nK} \left[\sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + \sum_{k=1}^K E \left(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)} \right) \right] \\
&\quad + \frac{1}{K} \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) \\
&= \frac{2(K-1)}{nK} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + \frac{2K-n-2}{nK} \sum_{k=1}^K E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) - \frac{1}{K} \sum_{k=1}^K E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})
\end{aligned}$$

notice that:

$$\sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) = \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) + \sum_{k=1}^K E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})$$

Thus,

$$\begin{aligned}
& \widehat{\langle X, X \rangle}_T^{(TSRV, K)} - [Z, Z]_T^{(avg, K)} \\
&= 2 \cdot \frac{\sqrt{n}}{K} \underbrace{\left(\underbrace{\frac{K-1}{n}}_{o_p(1)} \cdot M_T^{(1)} + \underbrace{\frac{n-K+1}{n}}_{1+o_p(1)} \cdot M_T^{(2)} - M_T^{(3)} \right)}_{\text{Mixed Normal}} + O_p \left(\frac{1}{\sqrt{K}} \right) \\
&\quad + \underbrace{\frac{2(K-1)}{nK} \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} E \left(\epsilon_{t_i}^2 | \omega^{(0)} \right) - \frac{n-2K+2}{nK} \left[\sum_{k=1}^K E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + \sum_{k=1}^K E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right]}_{\text{Edge Effect in tradition TSRV}}
\end{aligned} \tag{53}$$

11.2 Proof of Lemma 2

Proof.

$$\begin{aligned}
[Y, Y]_T^{(avg, K)} &= [Z, Z]_T^{(avg, K)} + [\epsilon, \epsilon]_T^{(avg, K)} + O_p \left(\frac{1}{\sqrt{K}} \right) \\
&= [Z, Z]_T^{(avg, K)} + \frac{2\sqrt{n}}{K} \left(M_T^{(1)} - M_T^{(3)} \right) + \frac{2}{K} \sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) \\
&\quad + \frac{1}{K} \sum_{k=1}^K \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right) + O_p \left(\frac{1}{\sqrt{K}} \right)
\end{aligned}$$

Remember that $M_T^{(3)}$ is asymptotically normal with asymptotic variance $\frac{1}{T} \int_0^T (g(X_t))^2 dt$. Now, I claim that $M_T^{(1)}$ is also asymptotically normal.

Denote $h_t(\omega^{(0)}) \equiv E(\epsilon_{t_i}^4 | \omega^{(0)})$. Remember the definitions of $M_T^{(1)}$ and $M_T^{(3)}$: $M_T^{(3)} = \frac{1}{\sqrt{n}} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_i} \epsilon_{t_i, -}$ and $M_T^{(1)} = \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)}) \right]$. To apply the martingale central limit theorem on $M_T^{(1)}$ and $M_T^{(3)}$, let we calculate the relevant predictable quadratic variance/covariance.

$$\begin{aligned}
\langle M^{(1)}, M^{(1)} \rangle_T | \omega^{(0)} &= \frac{1}{n} \sum_{i=0}^n E \left[\left(\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \mathcal{F}_{t_{i-1}}^{(1)}) \right)^2 \right] | \omega^{(0)} \\
&= \frac{1}{n} \sum_{i=0}^n \left[E(\epsilon_{t_i}^4 | \mathcal{F}_{t_{i-1}}^{(1)}) - \left(E(\epsilon_{t_i}^2 | \mathcal{F}_{t_{i-1}}^{(1)}) \right)^2 \right] \cdot \frac{T}{n} | \omega^{(0)} \\
&= \frac{1}{T} \sum_{i=0}^n \left[h_{t_i}(\omega^{(0)}) - \left(g_{t_i}(\omega^{(0)}) \right)^2 \right] \cdot \frac{T}{n} \\
&\longrightarrow \frac{1}{T} \int_0^T h_t(\omega^{(0)}) - \left(g_t(\omega^{(0)}) \right)^2 dt
\end{aligned}$$

the positiveness of the predictable quadratic variance is guaranteed by the Jensen's inequality.

Furthermore, let's see the predictable quadratic covariance $\langle M^{(1)}, M^{(3)} \rangle_T | \omega^{(0)}$.

$$\begin{aligned}
\langle M^{(1)}, M^{(3)} \rangle_T | \omega^{(0)} &= \frac{1}{n} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} Cov \left(\epsilon_{t_i} \epsilon_{t_{i,-}}, \epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)}) | \mathcal{F}_{t_{i-1}}^{(1)} \right) | \omega^{(0)} \\
&= \frac{1}{n} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left[\left(\epsilon_{t_i} \epsilon_{t_{i,-}} - E(\epsilon_{t_i} \epsilon_{t_{i,-}} | \mathcal{F}_{t_{i-1}}^{(1)}) \right) \cdot \left(\epsilon_{t_i}^2 - E(\epsilon_{t_i}^2 | \omega^{(0)}) \right) | \mathcal{F}_{t_{i-1}}^{(1)} \right] | \omega^{(0)} \\
&= \frac{1}{n} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^3 \epsilon_{t_{i,-}} | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)} \right) - E \left(\epsilon_{t_i} \epsilon_{t_{i,-}} | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)} \right) \cdot E(\epsilon_{t_i}^2 | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)})
\end{aligned}$$

Note that

$$E \left(\epsilon_{t_i} \epsilon_{t_{i,-}} | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)} \right) = \epsilon_{t_{i,-}} \cdot E \left(\epsilon_{t_i} | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)} \right) = 0$$

so,

$$\begin{aligned}
\langle M^{(1)}, M^{(3)} \rangle_T | \omega^{(0)} &= \frac{1}{n} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E \left(\epsilon_{t_i}^3 \epsilon_{t_{i,-}} | \omega^{(0)}, \mathcal{F}_{t_{i-1}}^{(1)} \right) = \frac{1}{n} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_{i,-}} E \left(E(\epsilon_{t_i}^3 | \omega^{(0)}) | \mathcal{F}_{t_{i-1}}^{(1)} \right) \\
&= E \left(\left(\langle M^{(1)}, M^{(3)} \rangle_T \right)^2 \cdot I_{\{\tau_l > T\}} | \omega^{(0)} \right) \\
&= E \left(\frac{1}{n^2} \left(\sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} \epsilon_{t_{i,-}} E \left(E(\epsilon_{t_i}^3 | \omega^{(0)}) | \mathcal{F}_{t_{i-1}}^{(1)} \right) \right)^2 \cdot I_{\{\tau_l > T\}} | \omega^{(0)} \right) \\
&= \frac{1}{n^2} \sum_{k=1}^K \sum_{t_i \in \mathcal{G}^{(k)}} E(\epsilon_{t_{i,-}}^2 | \omega^{(0)}) \cdot \left(E \left(E(\epsilon_{t_i}^3 | \omega^{(0)}) | \mathcal{F}_{t_{i-1}}^{(1)} \right) \right)^2 \cdot I_{\{\tau_l > T\}} \\
&\leq \frac{n - K + 1}{n^2} \cdot M_{(3,l)}^2 \cdot M_{(2,l)} \cdot I_{\{\tau_l > T\}}
\end{aligned}$$

By **Proposition 1** and the fact that $\mathbb{P}\{\tau_l > T\} \longrightarrow 1$ as $l \longrightarrow \infty$, we know:

$$\langle M^{(1)}, M^{(3)} \rangle_T = O_p \left(\frac{1}{\sqrt{n}} \right) \longrightarrow^P 0$$

By the conditional Lyapunov condition, the limiting predictable quadratic variations of are $\frac{1}{T} \int_0^T \left(g_t(\omega^{(0)}) \right)^2 dt$ and $\frac{1}{T} \int_0^T h_t(\omega^{(0)}) - \left(g_t(\omega^{(0)}) \right)^2 dt$, and the limiting predictable covariation tending to zero, these limiting quantities are all measurable in (the completions of) all the σ -fields $\{\mathcal{F}_{t_i}\}_{i=0}^n$. By martingale central limit theorem,

$$M_T^{(3)} \xrightarrow{\mathbb{P}} \mathcal{MN} \left(0, \frac{1}{T} \int_0^T h_t(\omega^{(0)}) - \left(g_t(\omega^{(0)}) \right)^2 dt \right)$$

Since

$$M_T^{(1)} \xrightarrow{\mathbb{P}} \mathcal{MN} \left(0, \frac{1}{T} \int_0^T \left(g_t(\omega^{(0)}) \right)^2 dt \right)$$

and $M_T^{(1)}$ and $M_T^{(3)}$ are asymptotically independent, we have:

$$M_T^{(1)} - M_T^{(3)} \xrightarrow{\mathbb{P}} \mathcal{MN}\left(0, \frac{1}{T} \int_0^T h_t(\omega^{(0)}) dt\right)$$

Thus, by setting $\frac{\sqrt{n}}{K} \rightarrow 0$, we have

$$\begin{aligned} & [Y, Y]_T^{(avg, K)} - [Z, Z]_T^{(avg, K)} \\ &= \underbrace{\sum_{k=1}^K \sum_{t_i \in \tilde{\mathcal{G}}^{(k)}} \frac{2}{K} E(\epsilon_{t_i}^2 | \omega^{(0)}) + \sum_{k=1}^K \frac{1}{K} \left(E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}) \right)}_{\text{bias of } [Y, Y]_T^{(avg, K)}} + o_p(1) \end{aligned}$$

□

11.3 Proof of Theorem 1

Proof. To mathematically formulate this idea of the proof, we need to introduce some notations first:

$$\begin{aligned} \mathcal{G}^{(\min)} &= \left\{ \min \mathcal{G}^{(1)}, \min \mathcal{G}^{(2)}, \dots, \min \mathcal{G}^{(K)} \right\} \\ \mathcal{G}^{(\max)} &= \left\{ \max \mathcal{G}^{(1)}, \max \mathcal{G}^{(2)}, \dots, \max \mathcal{G}^{(K)} \right\} \end{aligned}$$

Thus, we can describe the original time points as:

$$\{t_0, t_1, t_2, \dots, t_n\} = \mathcal{G}^{(\min)} \cup \left(\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)} \right) \cup \mathcal{G}^{(\max)}$$

Also, we have $\mathcal{G}_k^{(\min)} = \min \mathcal{G}^{(k)}$ and $\mathcal{G}_k^{(\max)} = \max \mathcal{G}^{(k)}$ for $k = 1, 2, \dots, K$, and $|\mathcal{G}^{(\min)}| = K$, $|\mathcal{G}^{(\max)}| = K$. Beside, define $\mathcal{G}_{K+1}^{(\min)}$ as the right immediate neighbor of $\max \mathcal{G}^{(\min)} = \min \mathcal{G}^{(K)}$ in the full grid \mathcal{G} , and define $\mathcal{G}_0^{(\max)}$ as the left immediate neighbor of $\min \mathcal{G}^{(\max)} = \mathcal{G}_1^{(\max)}$ in the full grid \mathcal{G} . Furthermore, we define $[Y, Y]_{\mathcal{H}}$ as the realized variance of process $\{Y_t\}$ computed from the fastest time scale on the grid \mathcal{H} .

Now, the new version of realized variance in [Kalnina and Linton \[2008\]](#) can be written as:

$$\widetilde{[Y, Y]_T^{\{n\}}} = \frac{1}{2} [Y, Y]_{\mathcal{G}^{(\min)}} + [Y, Y]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + \frac{1}{2} [Y, Y]_{\mathcal{G}^{(\max)}}$$

Since for any grid \mathcal{H} , $[Y, Y]_{\mathcal{H}} = [Z, Z]_{\mathcal{H}} + 2[Z, \epsilon]_{\mathcal{H}} + [\epsilon, \epsilon]_{\mathcal{H}}$, we have:

$$\begin{aligned} [Y, Y]_T^{\{n\}} &= \frac{1}{2} ([Z, Z]_{\mathcal{G}^{(\min)}} + 2[Z, \epsilon]_{\mathcal{G}^{(\min)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(\min)}}) + \frac{1}{2} ([Z, Z]_{\mathcal{G}^{(\max)}} + 2[Z, \epsilon]_{\mathcal{G}^{(\max)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(\max)}}) \\ &\quad + \left([Z, Z]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + 2[Z, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + [\epsilon, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} \right) \\ &= \left([Z, \epsilon]_{\mathcal{G}^{(\min)}} + 2[Z, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}} \right) + \frac{1}{2} [\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} + [\epsilon, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + \frac{1}{2} [\epsilon, \epsilon]_{\mathcal{G}^{(\max)}} + O_p(1) \end{aligned}$$

Note that

$$\begin{aligned} & [Z, \epsilon]_{\mathcal{G}^{(\min)}} + 2[Z, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}} \\ &= 2 \left([Z, \epsilon]_{\mathcal{G}^{(\min)}} + [Z, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}} \right) - ([Z, \epsilon]_{\mathcal{G}^{(\min)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}}) \\ &= 2[Z, \epsilon]_{\mathcal{G}^{(\min)} \cup (\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}) \cup \mathcal{G}^{(\max)}} - ([Z, \epsilon]_{\mathcal{G}^{(\min)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}}) \\ &= 2[Z, \epsilon]_T^{(all)} - ([Z, \epsilon]_{\mathcal{G}^{(\min)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}}) \end{aligned}$$

Thus,

$$[Z, \epsilon]_T^{(all)} \leq [Z, \epsilon]_{\mathcal{G}^{(\min)}} + 2[Z, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + [Z, \epsilon]_{\mathcal{G}^{(\max)}} \leq 2[Z, \epsilon]_T^{(all)}$$

Define $\Delta Z_{t_i} = Z_{t_i} - Z_{t_{i-1}}$, then

$$\begin{aligned}
E \left(\left([Z, \epsilon]_T^{(all)} \right)^2 I_{\{\tau_l > T\}} | \omega^{(0)} \right) &= E \left(\left(\sum_{i=1}^n \Delta Z_{t_i} (\epsilon_{t_i} - \epsilon_{t_{i-1}}) \right)^2 I_{\{\tau_l > T\}} | \omega^{(0)} \right) \\
&= I_{\{\tau_l > T\}} E \left(\sum_{i=1}^n \sum_{j=1}^n \Delta Z_{t_i} \Delta Z_{t_j} (\epsilon_{t_i} - \epsilon_{t_{i-1}}) (\epsilon_{t_j} - \epsilon_{t_{j-1}}) | \omega^{(0)} \right) \\
&= I_{\{\tau_l > T\}} \sum_{i=1}^n \sum_{j=1}^n \Delta Z_{t_i} \Delta Z_{t_j} E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}}) (\epsilon_{t_j} - \epsilon_{t_{j-1}}) | \omega^{(0)} \right]
\end{aligned}$$

Note that since $\epsilon_{t_i} | X = Y_{t_i} | X - f(X_{t_i})$, by the assumption that $\{Y_{t_i} | \omega^{(0)}\}_{i=0}^n$ are independent, the noises are mutually independent conditioning on the whole path of latent process X , thus

$$E[(\epsilon_{t_i} - \epsilon_{t_{i-1}})(\epsilon_{t_j} - \epsilon_{t_{j-1}}) | \omega^{(0)}] = \begin{cases} -E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}), & j = i - 1 \\ E(\epsilon_{t_{i-1}}^2) + \tilde{E}(\epsilon_{t_i}^2 | \omega^{(0)}), & j = i \\ -E(\epsilon_{t_i}^2 | \omega^{(0)}), & j = i + 1 \\ 0, & \text{otherwise} \end{cases}$$

So, if $\tau_l > T$, we have:

$$\begin{aligned}
&\sum_{i=1}^n \sum_{j=1}^n \Delta Z_{t_i} \Delta Z_{t_j} E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}}) (\epsilon_{t_j} - \epsilon_{t_{j-1}}) | \omega^{(0)} \right] \\
&= \sum_{i=1}^n (\Delta Z_{t_i})^2 \left(E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) + E(\epsilon_{t_i}^2 | \omega^{(0)}) \right) - \sum_{i=1}^n \left[\Delta Z_{t_{i-1}} \Delta Z_{t_i} E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) I_{\{2 \leq i \leq n\}} + \Delta Z_{t_i} \Delta Z_{t_{i+1}} E(\epsilon_{t_i}^2 | \omega^{(0)}) I_{\{1 \leq i \leq n-1\}} \right] \\
&= \sum_{i=0}^{n-1} (\Delta Z_{t_{i+1}})^2 E(\epsilon_{t_i}^2 | \omega^{(0)}) + \sum_{i=1}^n (\Delta Z_{t_i})^2 E(\epsilon_{t_i}^2 | \omega^{(0)}) - 2 \sum_{i=1}^{n-1} \Delta Z_{t_i} \Delta Z_{t_{i+1}} E(\epsilon_{t_i}^2 | \omega^{(0)}) \\
&= \sum_{i=1}^{n-1} \left((\Delta Z_{t_{i+1}})^2 + (\Delta Z_{t_i})^2 - 2 \Delta Z_{t_i} \Delta Z_{t_{i+1}} \right) E(\epsilon_{t_i}^2 | \omega^{(0)}) + (\Delta Z_{t_1})^2 E(\epsilon_{t_0}^2 | \omega^{(0)}) + (\Delta Z_{t_n})^2 E(\epsilon_{t_n}^2 | \omega^{(0)}) \\
&\leq M_{(2,l)} \cdot \left[\sum_{i=1}^{n-1} (\Delta Z_{t_{i+1}})^2 + \sum_{i=1}^{n-1} (\Delta Z_{t_i})^2 - 2 \sum_{i=1}^{n-1} \Delta Z_{t_i} \Delta Z_{t_{i+1}} + (\Delta Z_{t_1})^2 + (\Delta Z_{t_n})^2 \right] \\
&\leq 2M_{(2,l)} \cdot \left[\sum_{i=1}^n (\Delta Z_{t_i})^2 - \sum_{i=1}^{n-1} \Delta Z_{t_i} \Delta Z_{t_{i+1}} \right] \leq 2M_{(2,l)} \cdot \left([Z, Z]_T^{(all)} + \frac{1}{2} \sum_{i=1}^{n-1} [(\Delta Z_{t_i})^2 + (\Delta Z_{t_{i+1}})^2] \right) \\
&\leq 4M_{(2,l)} \cdot [Z, Z]_T^{(all)} = O_p(1)
\end{aligned}$$

By **Proposition 1** and $\mathbb{P}^{(0)}\{\tau_l > T\} \rightarrow 1$ as $l \rightarrow \infty$, we know $[Z, \epsilon]_T^{(all)} = O_p(1)$. So the following relation holds:

$$\begin{aligned}
\widetilde{[Y, Y]_T^{(all)}} &= \frac{1}{2} [\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} + [\epsilon, \epsilon]_{\bigcup_{k=1}^K \tilde{\mathcal{G}}^{(k)}} + \frac{1}{2} [\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} + O_p(1) \\
&= [\epsilon, \epsilon]_T^{(all)} - \frac{1}{2} ([\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(\max)}}) + O_p(1)
\end{aligned}$$

As in (49), We have already known:

$$[\epsilon, \epsilon]_T^{(all)} = 2\sqrt{n} \left(M_T^{(1)} - M_T^{(2)} \right) + 2 \sum_{i=0}^n E(\epsilon_{t_i}^2 | \omega^{(0)}) + O_p(1)$$

Define the following quantities:

$$\underline{m}_T^{(1)} \equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right] \quad (54)$$

$$\underline{m}_T^{(2)} \equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \epsilon_{\mathcal{G}_{k+1}^{(\min)}} \epsilon_{\mathcal{G}_k^{(\min)}} \quad (55)$$

$$\bar{m}_T^{(1)} \equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\max)}}^2 - g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) \right] \quad (56)$$

$$\bar{m}_T^{(2)} \equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\max)}} \epsilon_{\mathcal{G}_{k-1}^{(\max)}} \quad (57)$$

Actually, $\underline{m}_T^{(1)}$ and $\bar{m}_T^{(1)}$ are the “tiny” version of $M_T^{(1)}$, more specifically, there are constructed by the same way, but just use the first K realizations and the last K realizations of the noise process, respectively. So, $\underline{m}_T^{(1)}$ and $\bar{m}_T^{(1)}$ are asymptotically normal, with asymptotic variances smaller than the asymptotic variance of $M_T^{(1)}$. It is also true for $\underline{m}_T^{(2)}$ and $\bar{m}_T^{(2)}$.

Then,

$$\begin{aligned} [\epsilon, \epsilon]_{\mathcal{G}^{(\min)}} &= \sum_{k=1}^K \left(\epsilon_{\mathcal{G}_{k+1}^{(\min)}} - \epsilon_{\mathcal{G}_k^{(\min)}} \right)^2 = \sum_{k=1}^K \left(\epsilon_{\mathcal{G}_{k+1}^{(\min)}}^2 + \epsilon_{\mathcal{G}_k^{(\min)}}^2 - 2\epsilon_{\mathcal{G}_{k+1}^{(\min)}} \epsilon_{\mathcal{G}_k^{(\min)}} \right) \\ &= 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\min)}}^2 - 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_{k+1}^{(\min)}} \epsilon_{\mathcal{G}_k^{(\min)}} + \underbrace{\left(\epsilon_{\mathcal{G}_{K+1}^{(\min)}}^2 - \epsilon_{\mathcal{G}_1^{(\min)}}^2 \right)}_{O_p(1)} \\ &= 2 \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\min)}}^2 - E \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 | \omega^{(0)} \right) \right] - 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_{k+1}^{(\min)}} \epsilon_{\mathcal{G}_k^{(\min)}} + 2 \sum_{k=1}^K E \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 | X \right) + O_p(1) \\ &= 2\sqrt{K} \left(\underline{m}_T^{(1)} - \underline{m}_T^{(2)} \right) + 2 \sum_{k=1}^K E \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 | \omega^{(0)} \right) + O_p(1) \end{aligned}$$

$$\begin{aligned} [\epsilon, \epsilon]_{\mathcal{G}^{(\max)}} &= \sum_{k=0}^{K-1} \left(\epsilon_{\mathcal{G}_{k+1}^{(\max)}} - \epsilon_{\mathcal{G}_k^{(\max)}} \right)^2 = \sum_{k=0}^{K-1} \left(\epsilon_{\mathcal{G}_{k+1}^{(\max)}}^2 + \epsilon_{\mathcal{G}_k^{(\max)}}^2 - 2\epsilon_{\mathcal{G}_{k+1}^{(\max)}} \epsilon_{\mathcal{G}_k^{(\max)}} \right) \\ &= 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\max)}}^2 - 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\max)}} \epsilon_{\mathcal{G}_{k-1}^{(\max)}} + \underbrace{\left(\epsilon_{\mathcal{G}_0^{(\max)}}^2 - \epsilon_{\mathcal{G}_K^{(\max)}}^2 \right)}_{O_p(1)} \\ &= 2 \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(\max)}}^2 - E \left(\epsilon_{\mathcal{G}_k^{(\max)}}^2 | X \right) \right] - 2 \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(\max)}} \epsilon_{\mathcal{G}_{k-1}^{(\max)}} + 2 \sum_{k=1}^K E \left(\epsilon_{\mathcal{G}_k^{(\max)}}^2 | \omega^{(0)} \right) + O_p(1) \\ &= 2\sqrt{K} \left(\bar{m}_T^{(1)} - \bar{m}_T^{(2)} \right) + 2 \sum_{k=1}^K E \left(\epsilon_{\mathcal{G}_k^{(\max)}}^2 | \omega^{(0)} \right) + O_p(1) \end{aligned}$$

Therefore, the difference between sample-weighted TSRV and the averaged realized variance based on the theoretical quantity Z is:

$$\begin{aligned}
& \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - [Z, Z]_T^{(avg, K)} \\
&= [Y, Y]_T^{(avg, K)} - [Z, Z]_T^{(avg, K)} - \frac{1}{K} \widehat{[Y, Y]}_T^{(all)} \\
&= [\epsilon, \epsilon]_T^{(avg, K)} - \frac{1}{K} \left([\epsilon, \epsilon]_T^{(all)} - \frac{1}{2} ([\epsilon, \epsilon]_{\mathcal{G}^{(min)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(max)}}) + O_p(1) \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \\
&= \underbrace{\frac{2\sqrt{n}}{K} (M_T^{(1)} - M_T^{(3)}) + \frac{2}{K} \sum_{k=1}^K \sum_{t_i \in \hat{\mathcal{G}}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) + \frac{1}{K} \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)}))}_{[\epsilon, \epsilon]_T^{(avg)}} \\
&\quad - \underbrace{\left(\frac{2\sqrt{n}}{K} (M_T^{(1)} - M_T^{(2)}) + \frac{2}{K} \left[\sum_{k=1}^K \sum_{t_i \in \hat{\mathcal{G}}^{(k)}} E(\epsilon_{t_i}^2 | \omega^{(0)}) + \sum_{k=1}^K (E(\epsilon_{\min \mathcal{G}^{(k)}}^2 | \omega^{(0)}) + E(\epsilon_{\max \mathcal{G}^{(k)}}^2 | \omega^{(0)})) \right] \right)}_{\frac{1}{K} [\epsilon, \epsilon]_T^{(all)}} \\
&\quad + \underbrace{\frac{1}{\sqrt{K}} (\underline{m}_T^{(1)} - \underline{m}_T^{(2)}) + \frac{1}{K} \sum_{k=1}^K E(\epsilon_{\mathcal{G}_k^{(min)}}^2 | \omega^{(0)}) + \frac{1}{\sqrt{K}} (\bar{m}_T^{(1)} - \bar{m}_T^{(2)}) + \frac{1}{K} \sum_{k=1}^K E(\epsilon_{\mathcal{G}_k^{(max)}}^2 | \omega^{(0)})}_{\frac{1}{2K} ([\epsilon, \epsilon]_{\mathcal{G}^{(min)}} + [\epsilon, \epsilon]_{\mathcal{G}^{(max)}})} + O_p \left(\frac{1}{\sqrt{K}} \right)
\end{aligned} \tag{58}$$

$$\tag{59}$$

So,

$$\begin{aligned}
& \widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - [Z, Z]_T^{(avg, K)} \\
&= \frac{2\sqrt{n}}{K} (M_T^{(2)} - M_T^{(3)}) + \frac{1}{\sqrt{K}} (\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)}) + O_p \left(\frac{1}{\sqrt{K}} \right)
\end{aligned} \tag{60}$$

$$\begin{aligned}
& \text{So, } \frac{K}{\sqrt{n}} \left(\widehat{\langle X, X \rangle}_T^{(SW-TSRV, K)} - [Z, Z]_T^{(avg, K)} \right) \\
&= 2(M_T^{(2)} - M_T^{(3)}) + \sqrt{\frac{K}{n}} (\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)}) + O_p \left(\frac{1}{\sqrt{K}} \right) \\
&= 2(M_T^{(2)} - M_T^{(3)}) + o_p(1) \\
&\xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \frac{8}{T} \int_0^T (g_t(\omega^{(0)}))^2 dt \right)
\end{aligned}$$

For the remaining argument discussing the error term due to discretization, i.e., $[Z, Z]_T^{(avg)} - \langle Z, Z \rangle_T$. We can combine the result of **Lemma ??**, and get the claim the **Theorem 1**. \square

11.4 Proof of Theorem 2

Proof. To prove the limiting distribution is normal, again, we will use martingale limit central by exploiting the discrete predictable quadratic variation.

$$\begin{aligned}
\langle \underline{m}^{(2)}, \underline{m}^{(2)} \rangle_T | \omega^{(0)} &= \frac{1}{K} \sum_{k=1}^K E \left(\epsilon_{\mathcal{G}_k^{(min)}}^2 \epsilon_{\mathcal{G}_{k+1}^{(min)}}^2 | \mathcal{F}_{\mathcal{G}_k^{(min)}}^{(1)} \right) | \omega^{(0)} = \frac{1}{K} \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(min)}}^2 E \left(\epsilon_{\mathcal{G}_{k+1}^{(min)}}^2 | \mathcal{F}_{\mathcal{G}_k^{(min)}}^{(1)} \right) | \omega^{(0)} \\
&= \frac{1}{K} \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(min)}}^2 E \left(\epsilon_{\mathcal{G}_{k+1}^{(min)}}^2 | \mathcal{F}_{\mathcal{G}_k^{(min)}}^{(1)}, \omega^{(0)} \right) = \frac{1}{K} \sum_{k=1}^K \epsilon_{\mathcal{G}_k^{(min)}}^2 g_{\mathcal{G}_{k+1}^{(min)}}(\omega^{(0)}) \\
&= \frac{1}{K} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(min)}}^2 - g_{\mathcal{G}_k^{(min)}}(\omega^{(0)}) + g_{\mathcal{G}_k^{(min)}}(\omega^{(0)}) \right] g_{\mathcal{G}_{k+1}^{(min)}}(\omega^{(0)}) \\
&= \frac{1}{K} \sum_{k=1}^K \left[\epsilon_{\mathcal{G}_k^{(min)}}^2 - g_{\mathcal{G}_k^{(min)}}(\omega^{(0)}) \right] g_{\mathcal{G}_{k+1}^{(min)}}(\omega^{(0)}) + \frac{1}{K} \sum_{k=1}^K g_{\mathcal{G}_k^{(min)}}(\omega^{(0)}) g_{\mathcal{G}_{k+1}^{(min)}}(\omega^{(0)})
\end{aligned}$$

Since $K/n \rightarrow 0$, $\mathcal{G}_K^{(\min)} \rightarrow 0$, so $\frac{1}{K} \sum_{k=1}^K g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) \rightarrow (g_0(\omega^{(0)}))^2$
 Besides,

$$\begin{aligned}
 E \left(\left(\frac{1}{K} \sum_{k=1}^K \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right) g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) \mathbf{1}_{\{\tau_l > T\}} \right)^2 \middle| \omega^{(0)} \right) \\
 &= \frac{1}{K^2} \sum_{k=1}^K E \left(\left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2 \left(g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) \right)^2 \middle| \omega^{(0)} \right) \mathbf{1}_{\{\tau_l > T\}} \\
 &= \frac{1}{K^2} \sum_{k=1}^K E \left(\left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2 \middle| \omega^{(0)} \right) \left(g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) \right)^2 \mathbf{1}_{\{\tau_l > T\}} \\
 &\leq \frac{1}{K^2} \sum_{k=1}^K M_{(4,l)} \cdot M_{(2,l)}^2 = O_p \left(\frac{1}{K} \right)
 \end{aligned}$$

By **Proposition 1** and the fact that $P(\tau_l > T) \xrightarrow{\mathbb{P}} 1$ as $l \rightarrow \infty$, we know

$$\frac{1}{K} \sum_{k=1}^K \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right) g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) \xrightarrow{\mathbb{P}} 0$$

So,

$$\langle \underline{\mathbf{m}}^{(2)}, \underline{\mathbf{m}}^{(2)} \rangle_T | X = \frac{1}{K} \sum_{k=1}^K g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) g_{\mathcal{G}_{k+1}^{(\min)}}(\omega^{(0)}) + o_p(1) \rightarrow (g_0(\omega^{(0)}))^2 \quad (61)$$

By almost the same calculation, we can get

$$\langle \bar{\mathbf{m}}^{(2)}, \bar{\mathbf{m}}^{(2)} \rangle_T | X = \frac{1}{K} \sum_{k=1}^K g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) + o_p(1) \rightarrow (g_T(\omega^{(0)}))^2 \quad (62)$$

By martingale central limit theorem, we know $\underline{\mathbf{m}}_T^{(2)}$ and $\bar{\mathbf{m}}_T^{(2)}$ are asymptotically mixed normal:

$$\underline{\mathbf{m}}_T^{(2)} \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \left(g_0(\omega^{(0)}) \right)^2 \right) \quad (63)$$

$$\bar{\mathbf{m}}_T^{(2)} \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \left(g_T(\omega^{(0)}) \right)^2 \right) \quad (64)$$

Because $\underline{\mathbf{m}}_T^{(2)}$ and $\bar{\mathbf{m}}_T^{(2)}$ are constructed by noise on disjoint sets of observation times, so $\underline{\mathbf{m}}_T^{(2)}$ and $\bar{\mathbf{m}}_T^{(2)}$ are independent conditional on X , so

$$\left(\underline{\mathbf{m}}_T^{(2)} + \bar{\mathbf{m}}_T^{(2)} \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \left(g_0(\omega^{(0)}) \right)^2 + \left(g_T(\omega^{(0)}) \right)^2 \right)$$

Besides,

$$\begin{aligned}
 \langle \underline{\mathbf{m}}_T^{(1)}, \underline{\mathbf{m}}_T^{(1)} \rangle_T | \omega^{(0)} &= \frac{1}{K} \sum_{k=1}^K E \left[\left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2 \middle| \mathcal{F}_{\mathcal{G}_{k-1}^{(\min)}}^{(1)} \right] | \omega^{(0)} \\
 &= \frac{1}{K} \sum_{k=0}^K \left[E \left(\epsilon_{\mathcal{G}_k^{(\min)}}^4 \middle| \mathcal{F}_{\mathcal{G}_{k-1}^{(\min)}}^{(1)}, \omega^{(0)} \right) - \left(g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2 \right] \\
 &= \frac{1}{K} \sum_{i=0}^K \left[h_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) - \left(g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2 \right] \\
 &\rightarrow h_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) - \left(g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right)^2
 \end{aligned}$$

By the same calculation, we know

$$\langle \bar{\mathbf{m}}_T^{(1)}, \bar{\mathbf{m}}_T^{(1)} \rangle_T | X = \frac{1}{K} \sum_{i=0}^n \left[h_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) - \left(g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) \right)^2 \right] \rightarrow h_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) - \left(g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) \right)^2$$

Furthermore,

$$\begin{aligned}
\langle \underline{m}^{(1)}, \underline{m}^{(2)} \rangle_T | \omega^{(0)} &= \frac{1}{K} \sum_{k=1}^K Cov \left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}), \epsilon_{\mathcal{G}_k^{(\min)}} \epsilon_{\mathcal{G}_{k+1}^{(\min)}} | \mathcal{F}_{\mathcal{G}_k^{(\min)}}^{(1)} \right) | \omega^{(0)} \\
&= \frac{1}{K} \sum_{k=1}^K E \left[\left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right) \cdot \left(\epsilon_{\mathcal{G}_k^{(\min)}} \epsilon_{\mathcal{G}_{k+1}^{(\min)}} - E(\epsilon_{\mathcal{G}_k^{(\min)}} \epsilon_{\mathcal{G}_{k+1}^{(\min)}} | \mathcal{F}_{\mathcal{G}_k^{(\min)}}^{(1)}) \right) | \omega^{(0)}, \mathcal{F}_{\mathcal{G}_k^{(\min)}}^{(1)} \right] \\
&= \frac{1}{K} \sum_{k=1}^K \left[\left(\epsilon_{\mathcal{G}_k^{(\min)}}^2 - g_{\mathcal{G}_k^{(\min)}}(\omega^{(0)}) \right) \cdot \epsilon_{\mathcal{G}_k^{(\min)}} \underbrace{E \left(\epsilon_{\mathcal{G}_{k+1}^{(\min)}} - E(\epsilon_{\mathcal{G}_{k+1}^{(\min)}} | \mathcal{F}_{\mathcal{G}_k^{(\min)}}^{(1)}) \right) | \omega^{(0)}}_0 \right] \\
&= 0
\end{aligned}$$

and

$$\begin{aligned}
\langle \bar{m}^{(1)}, \bar{m}^{(2)} \rangle_T | \omega^{(0)} &= \frac{1}{K} \sum_{k=1}^K Cov \left(\epsilon_{\mathcal{G}_k^{(\max)}}^2 - g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}), \epsilon_{\mathcal{G}_{k-1}^{(\max)}} \epsilon_{\mathcal{G}_k^{(\max)}} | \mathcal{F}_{\mathcal{G}_{k-1}^{(\max)}}^{(1)} \right) | \omega^{(0)} \\
&= \frac{1}{K} \sum_{k=1}^K E \left[\left(\epsilon_{\mathcal{G}_k^{(\max)}}^2 - g_{\mathcal{G}_k^{(\max)}}(\omega^{(0)}) \right) \cdot \epsilon_{\mathcal{G}_{k-1}^{(\max)}} \left(\epsilon_{\mathcal{G}_k^{(\max)}} - E(\epsilon_{\mathcal{G}_k^{(\max)}} | \mathcal{F}_{\mathcal{G}_{k-1}^{(\max)}}^{(1)}) \right) | \omega^{(0)} \right] \\
&= \frac{1}{K} \sum_{k=1}^K \epsilon_{\mathcal{G}_{k-1}^{(\max)}} E \left(\epsilon_{\mathcal{G}_k^{(\max)}}^3 | \omega^{(0)}, \mathcal{F}_{\mathcal{G}_{k-1}^{(\max)}}^{(1)} \right)
\end{aligned}$$

$$\begin{aligned}
E \left(\left(\langle \bar{m}^{(1)}, \bar{m}^{(2)} \rangle_T \right)^2 \cdot I_{\{\tau_l > T\}} | \omega^{(0)} \right) &= E \left(\frac{1}{K^2} \left(\sum_{k=1}^K \epsilon_{\mathcal{G}_{k-1}^{(\max)}} E \left(\epsilon_{\mathcal{G}_k^{(\max)}}^3 | \omega^{(0)}, \mathcal{F}_{\mathcal{G}_{k-1}^{(\max)}}^{(1)} \right) \right)^2 \cdot I_{\{\tau_l > T\}} | \omega^{(0)} \right) \\
&= \frac{1}{K^2} \sum_{k=1}^K E(\epsilon_{\mathcal{G}_{k-1}^{(\max)}}^2 | \omega^{(0)}) \cdot \left(E \left(E(\epsilon_{\mathcal{G}_k^{(\max)}}^3 | \omega^{(0)}) | \mathcal{F}_{\mathcal{G}_{k-1}^{(\max)}}^{(1)} \right) \right)^2 \cdot I_{\{\tau_l > T\}} \\
&\leq \frac{1}{K} \cdot M_{(2,l)} \cdot M_{(3,l)}^2 \cdot I_{\{\tau_l > T\}}
\end{aligned}$$

By **Proposition 1** and the fact that $\mathbb{P}\{\tau_l > T\} \rightarrow 1$ as $l \rightarrow \infty$, we know:

$$\langle \bar{m}^{(1)}, \bar{m}^{(2)} \rangle_T = O_p \left(\frac{1}{\sqrt{K}} \right) \xrightarrow{\mathbb{P}} 0$$

Thus, $\underline{m}_T^{(1)}$, $\underline{m}_T^{(2)}$, $\bar{m}_T^{(1)}$ and $\bar{m}_T^{(2)}$ are asymptotically independent and mixed normals, and

$$\underline{m}_T^{(1)} - \underline{m}_T^{(2)} + \bar{m}_T^{(1)} - \bar{m}_T^{(2)} \xrightarrow{\mathcal{L}-s} \mathcal{MN}(0, h_0(\omega^{(0)}) + h_T(\omega^{(0)}))$$

□

11.5 Proof of Lemma 3

Proof.

$$\begin{aligned}
[Y, Y, Y, Y]_T^{(all)} &= \sum_{i=1}^n (Y_{t_i} - Y_{t_{i-1}})^4 = \sum_{i=1}^n [(Z_{t_i} - Z_{t_{i-1}}) + (\epsilon_{t_i} - \epsilon_{t_{i-1}})]^4 \\
&= \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^4 + \sum_{i=1}^n (Z_{t_i} - Z_{t_{i-1}})^4 + 4 \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^3 (Z_{t_i} - Z_{t_{i-1}}) \\
&\quad + 6 \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 + 4 \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}}) (Z_{t_i} - Z_{t_{i-1}})^3
\end{aligned}$$

Since $f(X_t)$ is an Itô semimartingale, $\sum_{i=1}^n (Z_{t_i} - Z_{t_{i-1}})^4 = o_p(1)$. Since we have proved $[f(X), \epsilon]_T^{(all)} = \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 = O_p(1)$, thus

$$\begin{aligned}
& 6 \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 = O_p(1) \\
& E \left[\left(\sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^3 (Z_{t_i} - Z_{t_{i-1}}) \right)^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)} \right] \\
&= \sum_{i=1}^n E \left[\left((\epsilon_{t_i} - \epsilon_{t_{i-1}})^3 (Z_{t_i} - Z_{t_{i-1}}) \right)^2 | \omega^{(0)} \right] \\
&= \sum_{i=1}^n E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}})^4 \cdot (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 | \omega^{(0)} \right] \mathbf{1}_{\{\tau_l > T\}} \\
&\leq \left(2M_{(4,l)} + 6M_{(2,l)}^2 \right) \mathbf{1}_{\{\tau_l > T\}} \sum_{i=1}^n E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 | \omega^{(0)} \right] \\
&= O_p(1) \\
& E \left[\left(\sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}}) (Z_{t_i} - Z_{t_{i-1}})^3 \right)^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)} \right] \\
&= \sum_{i=1}^n E \left[\left((\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}}) \right)^6 | \omega^{(0)} \right] \\
&= \sum_{i=1}^n E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 \cdot (Z_{t_i} - Z_{t_{i-1}})^4 | \omega^{(0)} \right] \mathbf{1}_{\{\tau_l > T\}} \\
&\leq \sum_{i=1}^n E \left[(\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 | \omega^{(0)} \right] \mathbf{1}_{\{\tau_l > T\}} \cdot \sum_{i=1}^n (Z_{t_i} - Z_{t_{i-1}})^4 \\
&= O_p(1) \cdot o_p(1) = o_p(1)
\end{aligned}$$

So we know

$$\begin{aligned}
\sum_{i=1}^n (Z_{t_i} - Z_{t_{i-1}})^4 &= o_p(1) \\
\sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}}) (Z_{t_i} - Z_{t_{i-1}})^3 &= o_p(1) \\
\sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^2 (Z_{t_i} - Z_{t_{i-1}})^2 &= O_p(1) \\
\sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^3 (Z_{t_i} - Z_{t_{i-1}}) &= O_p(1)
\end{aligned}$$

thus,

$$\begin{aligned}
[Y, Y, Y, Y]_T^{(all)} &= \sum_{i=1}^n (\epsilon_{t_i} - \epsilon_{t_{i-1}})^4 + O_p(1) \\
&\equiv [\epsilon, \epsilon, \epsilon, \epsilon]_T^{(all)} + O_p(1)
\end{aligned} \tag{65}$$

Note that

$$\begin{aligned}
[\epsilon, \epsilon, \epsilon, \epsilon]_T^{(all)} &= \sum_{i=1}^n \left(\epsilon_{t_{i-1}}^4 - 4\epsilon_{t_{i-1}}^3 \epsilon_{t_i} + 6\epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 - 4\epsilon_{t_{i-1}} \epsilon_{t_i}^3 + \epsilon_{t_i}^4 \right) \\
&= 2 \sum_{i=1}^n \epsilon_{t_i}^4 + 6 \sum_{i=1}^n \epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 - 4 \sum_{i=1}^n \epsilon_{t_{i-1}}^3 \epsilon_{t_i} - 4 \sum_{i=1}^n \epsilon_{t_{i-1}} \epsilon_{t_i}^3 + (\epsilon_{t_0} - \epsilon_{t_n}) \\
&= 2 \sum_{i=1}^n \left[\epsilon_{t_i}^4 - E(\epsilon_{t_i}^4 | \omega^{(0)}) \right] + 6 \sum_{i=1}^n \left[\epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 - E(\epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 | \omega^{(0)}) \right] - 4 \sum_{i=1}^n \epsilon_{t_{i-1}}^3 \epsilon_{t_i} \\
&\quad - 4 \sum_{i=1}^n \epsilon_{t_{i-1}} \epsilon_{t_i}^3 + 2 \sum_{i=1}^n E(\epsilon_{t_i}^4 | \omega^{(0)}) + 6 \sum_{i=1}^n E(\epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 | \omega^{(0)}) + (\epsilon_{t_0} - \epsilon_{t_n})
\end{aligned}$$

Define

$$\begin{aligned}
L_T^{(1)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\epsilon_{t_i}^4 - E(\epsilon_{t_i}^4 | \omega^{(0)}) \right] \\
L_T^{(2)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \left[\epsilon_{t_{i-1}}^2 \epsilon_{t_i}^2 - E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) E(\epsilon_{t_i}^2 | \omega^{(0)}) \right] \\
L_T^{(3)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \epsilon_{t_{i-1}}^3 \epsilon_{t_i} \\
L_T^{(4)} &= \frac{1}{\sqrt{n}} \sum_{i=1}^n \epsilon_{t_{i-1}}^3 \epsilon_{t_i}^3
\end{aligned}$$

then we have

$$\begin{aligned}
[\epsilon, \epsilon, \epsilon, \epsilon]_T^{(all)} &= 2 \sum_{i=1}^n E(\epsilon_{t_i}^4 | \omega^{(0)}) + 6 \sum_{i=1}^n E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) E(\epsilon_{t_i}^2 | \omega^{(0)}) \\
&\quad + \sqrt{n} \left(2L_T^{(1)} + 6L_T^{(2)} - 4L_T^{(3)} - 4L_T^{(4)} \right) + O_p(1)
\end{aligned} \tag{66}$$

We can show that $L_T^{(1)}$, $L_T^{(2)}$, $L_T^{(3)}$ and $L_T^{(4)}$ are mixed normals. And observe that

$$\begin{aligned}
\frac{1}{n} \sum_{i=1}^n E(\epsilon_{t_i}^4 | \omega^{(0)}) &= \frac{1}{T} \sum_{i=1}^n E(\epsilon_{t_i}^4 | \omega^{(0)}) \frac{T}{n} \longrightarrow \frac{1}{T} \int_0^T h_t(\omega^{(0)}) dt \\
\frac{1}{n} \sum_{i=1}^n E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) E(\epsilon_{t_i}^2 | \omega^{(0)}) &= \frac{1}{T} \sum_{i=1}^n E(\epsilon_{t_{i-1}}^2 | \omega^{(0)}) E(\epsilon_{t_i}^2 | \omega^{(0)}) \frac{T}{n} \longrightarrow \frac{1}{T} \int_0^T g_t^2(\omega^{(0)}) dt
\end{aligned}$$

then the relation (24) follows. □

11.6 Proof of Lemma 4

Proof. Remember that $m_i = m_i^{(1)} - m_i^{(2)}$, so

$$\begin{aligned}
m_i^2 &= \left(m_i^{(1)} \right)^2 + \left(m_i^{(2)} \right)^2 - 2m_i^{(1)} m_i^{(2)} \\
m_i^4 &= \left(m_i^{(1)} \right)^4 - 4 \left(m_i^{(1)} \right)^3 m_i^{(2)} + 6 \left(m_i^{(1)} \right)^2 \left(m_i^{(2)} \right)^2 - 4m_i^{(1)} \left(m_i^{(2)} \right)^3 + \left(m_i^{(2)} \right)^4
\end{aligned}$$

For the ease of notion, let us denote $\epsilon_{(i-1)K_n+k}$ by $\xi_{i,k}$ for each $i \in \{1, 2, \dots, r_n\}$ and $k \in \{0, 1, 2, \dots, K\}$.. Note that under our new notation

$$\begin{aligned}
m_i^{(1)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \\
m_i^{(2)} &\equiv \frac{1}{\sqrt{K}} \sum_{k=1}^K \xi_{i,k-1} \xi_{i,k}
\end{aligned}$$

thus

$$\begin{aligned}
E \left[\left(m_i^{(1)} \right)^2 | \omega^{(0)} \right] &= \frac{1}{K} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 | \omega^{(0)} \right] = \frac{1}{K} \sum_{k=1}^K E \left[\xi_{i,k}^4 - \left(E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 | \omega^{(0)} \right] \\
&= \frac{1}{K} \sum_{k=1}^K \left[E(\xi_{i,k}^4 | \omega^{(0)}) - \left(E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \right] = E(\epsilon^4 | \omega^{(0)}) - \left(E(\epsilon^2 | \omega^{(0)}) \right)^2
\end{aligned}$$

$$\begin{aligned}
E \left[\left(m_i^{(2)} \right)^2 | \omega^{(0)} \right] &= \frac{1}{K} E \left[\left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right)^2 | \omega^{(0)} \right] = \frac{1}{K} \sum_{k=1}^K E \left[\xi_{i,k-1}^2 \xi_{i,k}^2 | \omega^{(0)} \right] \\
&= \frac{1}{K} \sum_{k=1}^K E(\xi_{i,k-1}^2 | \omega^{(0)}) E(\xi_{i,k}^2 | \omega^{(0)}) = \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 \\
E \left[m_i^{(1)} m_i^{(2)} | \omega^{(0)} \right] &= \frac{1}{K} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right) | \omega^{(0)} \right] \\
&= \frac{1}{K} E \left[\sum_{k=1}^K \sum_{j=1}^K \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot (\xi_{i,j-1} \xi_{i,j}) | \omega^{(0)} \right] \\
&= \frac{1}{K} \sum_{k=1}^K \sum_{j=1}^K E \left[\left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \xi_{i,j-1} \cdot \xi_{i,j} | \omega^{(0)} \right] = 0
\end{aligned}$$

Thus $E(m_i^2 | \omega^{(0)}) = E(\epsilon^4 | \omega^{(0)})$.

In order to evaluate $E(m_i^4 | \omega^{(0)})$, we need to evaluate $E \left(\left(m_i^{(1)} \right)^4 | \omega^{(0)} \right)$, $E \left(\left(m_i^{(1)} \right)^3 m_i^{(2)} | \omega^{(0)} \right)$, $E \left(\left(m_i^{(1)} \right)^2 \left(m_i^{(2)} \right)^2 | \omega^{(0)} \right)$, $E \left(m_i^{(1)} \left(m_i^{(2)} \right)^3 | \omega^{(0)} \right)$ and $E \left(\left(m_i^{(2)} \right)^4 | \omega^{(0)} \right)$ respectively:

$$\begin{aligned}
E \left[\left(m_i^{(1)} \right)^4 | \omega^{(0)} \right] &= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^4 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} \left[\sum_{k=1}^K E \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^4 + 6 \sum_{k=1}^K \sum_{j \neq k} E \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 E \left(\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)}) \right)^2 \right] \\
&= 6 \left[E(\epsilon^4 | \omega^{(0)}) - \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 \right] + O_p \left(\frac{1}{K} \right)
\end{aligned}$$

$$\begin{aligned}
E \left[\left(m_i^{(1)} \right)^3 m_i^{(2)} | \omega^{(0)} \right] &= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^3 \cdot \left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right) | \omega^{(0)} \right] \\
&= \frac{3}{K^2} E \left[\left(\sum_{k=1}^K \sum_{j \neq k} (\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}))^2 (\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)})) \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1} \xi_{i,j} \right) | \omega^{(0)} \right] \\
&= \frac{3}{K^2} \sum_{k=2}^K E \left[(\xi_{i,k-1}^2 - E(\xi_{i,k-1}^2 | \omega^{(0)}))^2 \xi_{i,k-1} \cdot (\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)})) \xi_{i,k} | \omega^{(0)} \right] \\
&\quad + \frac{3}{K^2} \sum_{k=1}^{K-1} E \left[(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}))^2 \xi_{i,k} \cdot (\xi_{i,k+1}^2 - E(\xi_{i,k+1}^2 | \omega^{(0)})) \xi_{i,k+1} | \omega^{(0)} \right] \\
&= \frac{3}{K^2} \sum_{k=2}^K E \left[(\xi_{i,k-1}^5 - 2\xi_{i,k-1}^3 E(\xi_{i,k-1}^2 | \omega^{(0)})) \cdot \xi_{i,k}^3 | \omega^{(0)} \right] \\
&\quad + \frac{3}{K^2} \sum_{k=1}^{K-1} E \left[(\xi_{i,k}^5 - 2\xi_{i,k}^3 E(\xi_{i,k}^2 | \omega^{(0)})) \cdot \xi_{i,k+1}^3 | \omega^{(0)} \right] \\
&= O_p \left(\frac{1}{K} \right)
\end{aligned}$$

$$\begin{aligned}
E \left[\left(m_i^{(1)} \right)^2 \left(m_i^{(2)} \right)^2 | \omega^{(0)} \right] &= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \cdot \left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right)^2 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K (\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}))^2 \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1} \xi_{i,j} \right)^2 | \omega^{(0)} \right] \\
&\quad + \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \sum_{j \neq k} (\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)})) (\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)})) \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1} \xi_{i,j} \right)^2 | \omega^{(0)} \right]
\end{aligned}$$

$$\begin{aligned}
& \text{Note that } \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1} \xi_{i,j} \right)^2 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \right) \cdot \left(\sum_{j=1}^K \xi_{j-1}^2 \xi_{i,j}^2 \right) | \omega^{(0)} \right] \\
&= \frac{1}{K^2} E \left[\sum_{k=1}^K \sum_{j \neq k, k+1} \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \cdot \xi_{j-1}^2 \xi_{i,j}^2 \right] + \frac{1}{K^2} E \left[\sum_{k=1}^K \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \cdot \xi_{i,k-1}^2 \xi_{i,k}^2 \right] \\
&\quad + \frac{1}{K^2} E \left[\sum_{k=1}^{K-1} \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right)^2 \cdot \xi_{i,k}^2 \xi_{i,k+1}^2 \right] \\
&= \left[E(\epsilon^4 | \omega^{(0)}) - \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 \right] \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 + O_p \left(\frac{1}{K} \right)
\end{aligned}$$

$$\begin{aligned}
& \text{and } \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \sum_{j \neq k} \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \left(\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)}) \right) \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1} \xi_{i,j} \right)^2 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \sum_{j \neq k} \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \left(\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)}) \right) \right) \cdot \left(\sum_{j=1}^K \xi_{i,j-1}^2 \xi_{i,j}^2 \right) | \omega^{(0)} \right] \\
&\quad + \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \sum_{j \neq k} \left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \left(\xi_{i,j}^2 - E(\xi_{i,j}^2 | \omega^{(0)}) \right) \right) \cdot \left(\sum_{j=1}^{K-1} \xi_{i,j-1} \xi_{i,j}^2 \xi_{i,j+1} \right) | \omega^{(0)} \right] \\
&= \frac{1}{K^2} E \left[\left(\sum_{k=2}^K \left(\xi_{i,k-1}^4 - \xi_{i,k-1}^2 E(\xi_{i,k-1}^2 | \omega^{(0)}) \right) \left(\xi_{i,k}^4 - \xi_{i,k}^2 E(\xi_{i,k}^2 | \omega^{(0)}) \right) \right) | \omega^{(0)} \right] \\
&\quad + \frac{1}{K^2} E \left[\left(\sum_{k=2}^{K-1} \left(\xi_{i,k-1}^3 - \xi_{i,k-1} E(\xi_{i,k-1}^2 | \omega^{(0)}) \right) \cdot \xi_{i,k}^2 \cdot \left(\xi_{i,k+1}^3 - \xi_{i,k+1} E(\xi_{i,k+1}^2 | \omega^{(0)}) \right) \right) | \omega^{(0)} \right] \\
&= \frac{1}{K^2} \left[\sum_{k=2}^K E \left(\xi_{i,k-1}^3 | \omega^{(0)} \right) \cdot E \left(\xi_{i,k}^3 | \omega^{(0)} \right) + \sum_{k=2}^{K-1} E \left(\xi_{i,k-1}^3 | \omega^{(0)} \right) \cdot E \left(\xi_{i,k}^2 | \omega^{(0)} \right) \cdot E \left(\xi_{i,k+1}^3 | \omega^{(0)} \right) \right] = O_p \left(\frac{1}{K} \right)
\end{aligned}$$

so we have

$$\begin{aligned}
& E \left[\left(m_i^{(1)} \right)^2 \left(m_i^{(2)} \right)^2 | \omega^{(0)} \right] = \left[E(\epsilon^4 | \omega^{(0)}) - \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 \right] \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 + O_p \left(\frac{1}{K} \right) \\
& E \left[m_i^{(1)} \left(m_i^{(2)} \right)^3 | \omega^{(0)} \right] = \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right)^3 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} \sum_{k=1}^K E \left[\left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \xi_{i,k-1}^3 \xi_{i,k}^3 \right] + \frac{1}{K^2} \sum_{k=1}^{K-1} E \left[\left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \xi_{i,k}^3 \xi_{i,k+1}^3 \right] \\
&\quad + \frac{1}{K^2} \sum_{k=2}^K E \left[\left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot 3 \xi_{i,k-2}^2 \xi_{i,k-1}^2 \cdot \xi_{i,k-1} \xi_{i,k} \right] \\
&\quad + \frac{1}{K^2} \sum_{k=1}^{K-1} E \left[\left(\xi_{i,k}^2 - E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot 3 \xi_{i,k} \xi_{i,k+1} \cdot \xi_{i,k+1}^2 \xi_{i,k+2}^2 \right] \\
&= \frac{1}{K^2} \sum_{k=1}^K E \left[\left(\xi_{i,k}^5 - \xi_{i,k}^3 E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \xi_{i,k-1}^3 \right] + \frac{1}{K^2} \sum_{k=1}^{K-1} E \left[\left(\xi_{i,k}^5 - \xi_{i,k}^3 E(\xi_{i,k}^2 | \omega^{(0)}) \right) \cdot \xi_{i,k+1}^3 \right] \\
&\quad + \frac{1}{K^2} \sum_{k=2}^K 3 E \left[\xi_{i,k}^3 \xi_{i,k-2}^2 \xi_{i,k-1}^3 \right] + \frac{1}{K^2} \sum_{k=1}^{K-1} 3 E \left[\xi_{i,k}^3 \xi_{i,k+1}^3 \xi_{i,k+2}^2 \right] \\
&= O_p \left(\frac{1}{K} \right)
\end{aligned}$$

$$\begin{aligned}
E \left[\left(m_i^{(2)} \right)^4 | \omega^{(0)} \right] &= \frac{1}{K^2} E \left[\left(\sum_{k=1}^K \xi_{i,k-1} \xi_{i,k} \right)^4 | \omega^{(0)} \right] \\
&= \frac{1}{K^2} \sum_{k=1}^K E \left(\xi_{i,k-1}^4 \xi_{i,k}^4 | \omega^{(0)} \right) + \frac{6}{K^2} \sum_{k=1}^K \sum_{j \neq k} E \left(\xi_{i,k-1}^2 \xi_{i,k}^2 \xi_{i,j-1}^2 \xi_{i,j}^2 | \omega^{(0)} \right) \\
&= \frac{6}{K^2} \sum_{k=1}^K E \left(\xi_{i,k-1}^2 \xi_{i,k}^4 \xi_{i,k+1}^2 | \omega^{(0)} \right) + \frac{6}{K^2} \sum_{k=1}^K \sum_{j \neq k, k+1} E \left(\xi_{i,k-1}^2 \xi_{i,k}^2 \xi_{i,j-1}^2 \xi_{i,j}^2 | \omega^{(0)} \right) + O_p \left(\frac{1}{K} \right) \\
&= 6(E(\epsilon^2 | \omega^{(0)}))^4 + O_p \left(\frac{1}{K} \right)
\end{aligned}$$

Thus, from the above calculation, we have:

$$\begin{aligned}
E(m_i^4 | \omega^{(0)}) &= E \left[\left(m_i^{(1)} \right)^4 | \omega^{(0)} \right] + 6E \left[\left(m_i^{(1)} \right)^2 \left(m_i^{(2)} \right)^2 | \omega^{(0)} \right] + E \left[\left(m_i^{(2)} \right)^4 | \omega^{(0)} \right] + O_p \left(\frac{1}{K} \right) \\
&= 6 \left[\left(E(\epsilon^4 | \omega^{(0)}) \right)^2 - E(\epsilon^4 | \omega^{(0)}) \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 + \left(E(\epsilon^2 | \omega^{(0)}) \right)^4 \right] + O_p \left(\frac{1}{K} \right)
\end{aligned}$$

□

11.7 Proof of Theorem 3

Proof. Under the assumption of this theorem, we know $g_t(\omega^{(0)})$ is a constant, $g_t(\omega^{(0)}) \equiv E(\epsilon^2 | \omega^{(0)})$. By the proof of **Theorem 2**, we know under the null hypothesis:

$$m_i \xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, E(\epsilon^4 | \omega^{(0)}) \right) \quad (67)$$

By continuous mapping theorem, we can get:

$$m_i^2 \xrightarrow{\mathcal{L}-s} E(\epsilon^4 | \omega^{(0)}) \cdot \chi_1^2(0)$$

where $\chi_1^2(0)$ denotes a centered chi-square random variable with degree of freedom 1, and independent of $\mathcal{F}^{(1)}$.

Note that

$$\begin{aligned}
V(Y, K_n, 2)_T^n &= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left| N(Y, K_n)_{[T_{i-1}, T_{i-1} + s_n]} \right|^2 = \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} (m_i + m_{i+s_n-1})^2 + o_p(1) \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 + m_{i+s_n-1}^2 + 2m_i m_{i+s_n-1} \right) + o_p(1) \\
&= \frac{1}{r_n - s_n + 1} \left[\sum_{i=1}^{r_n - s_n + 1} m_i^2 + \sum_{i=1}^{r_n - s_n + 1} m_{i+s_n-1}^2 + 2 \sum_{i=1}^{r_n - s_n + 1} m_i m_{i+s_n-1} \right] + o_p(1)
\end{aligned}$$

Since the noise process is assumed to be stationary, m_i and m_{i+s_n-1} are independent mean-0 martingales conditioning on X . Suppose the second moment is denote by $E(m^2 | X)$. Thus, by law of large number, we have:

$$\begin{aligned}
\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} m_i^2 | \omega^{(0)} &\xrightarrow{\mathcal{L}-s} \text{plim} E(m^2 | \omega^{(0)}) = E(\epsilon^4 | \omega^{(0)}) \\
\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} m_{i+s_n-1}^2 | \omega^{(0)} &\xrightarrow{\mathcal{L}-s} \text{plim} E(m^2 | \omega^{(0)}) = E(\epsilon^4 | \omega^{(0)}) \\
\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} m_i m_{i+s_n-1} | \omega^{(0)} &\xrightarrow{\mathcal{L}-s} \text{plim} E(m_i m_{i+s_n-1} | \omega^{(0)}) = \text{plim} E(m_i | \omega^{(0)}) E(m_{i+s_n-1} | \omega^{(0)}) = 0
\end{aligned}$$

Since $m_i^2 = O_p(1)$ $s_n = o_p(r_n)$, we know $\frac{1}{r_n - s_n + 1} \sum_{i=r_n - s_n + 2}^{r_n} m_i^2 = o_p(1)$ and $\frac{1}{r_n - s_n + 1} \sum_{i=1}^{s_n-1} m_i^2 = o_p(1)$. Therefore,

$$\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left| N(Y, K_n)_{[T_{i-1}, T_{i-1} + s_n]} \right|^2 - 2E(\epsilon^4 | \omega^{(0)}) \xrightarrow{\mathbb{P}} 0 \quad (68)$$

We have found the probability limit of the test statistic $\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left| N(Y, K_n)_{[T_{i-1}, T_{i-1} + s_n]} \right|^2$ in (68). Next, we decompose the LHS of (68) to find its asymptotic distribution.

$$\begin{aligned}
& \sqrt{r_n - s_n + 1} \left(V(Y, K_n, 2)_T^n - 2E(\epsilon^4 | \omega^{(0)}) \right) \\
&= \sqrt{r_n - s_n + 1} \left[\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left| N(Y, K)_{[T_{i-1}, T_{i-1} + s_n]} \right|^2 - 2E(\epsilon^4 | \omega^{(0)}) \right] \\
&= \sqrt{r_n - s_n + 1} \left[\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 + m_{i+s_n-1}^2 + 2m_i m_{i+s_n-1} \right) - 2E(\epsilon^4 | \omega^{(0)}) \right] \\
&= \frac{1}{\sqrt{r_n - s_n + 1}} \left[\sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(\epsilon^4 | \omega^{(0)}) \right) + \sum_{i=1}^{r_n - s_n + 1} \left(m_{i-s_n+1}^2 - E(\epsilon^4 | X) \right) \right] + \frac{2}{\sqrt{r_n - s_n + 1}} \sum_{i=1}^{r_n - s_n + 1} m_i m_{i-s_n+1} \\
&= 2 \left[\frac{1}{\sqrt{r_n - s_n + 1}} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) + \frac{1}{\sqrt{r_n - s_n + 1}} \sum_{i=1}^{r_n - s_n + 1} m_i m_{i-s_n+1} \right] \\
&\quad + \frac{1}{\sqrt{r_n - s_n + 1}} \left[\sum_{i=r_n - s_n + 2}^{r_n} \left(m_i^2 - E(\epsilon^4 | \omega^{(0)}) \right) - \sum_{i=1}^{s_n-1} \left(m_i^2 - E(\epsilon^4 | \omega^{(0)}) \right) \right]
\end{aligned} \tag{69}$$

the last equality hold since $E(m_i^2 | \omega^{(0)}) = E(\epsilon^4 | \omega^{(0)})$ by **Lemma 4**.

Define

$$\begin{aligned}
H_T^{(1)} &= \frac{1}{\sqrt{r_n - s_n + 1}} \sum_{i=1}^{r_n - s_n + 1} (m_i - E(m_i^2 | \omega^{(0)})) \\
H_T^{(2)} &= \frac{1}{\sqrt{r_n - s_n + 1}} \sum_{i=1}^{r_n - s_n + 1} m_i m_{i-s_n+1} \\
R_T &= \frac{1}{\sqrt{r_n - s_n + 1}} \left[\sum_{i=r_n - s_n + 2}^{r_n} (m_i^2 - E(\epsilon^4 | \omega^{(0)})) - \sum_{i=1}^{s_n-1} (m_i^2 - E(\epsilon^4 | \omega^{(0)})) \right]
\end{aligned}$$

then we have the following expression for the LHS of (68):

$$\sqrt{r_n - s_n + 1} \left(V(Y, K_n, 2)_T^n - 2E(\epsilon^4 | \omega^{(0)}) \right) = 2 \left(H_T^{(1)} + H_T^{(2)} \right) + R_T + o_p(1) \tag{70}$$

Furthermore, note that on the coarser filtered probability space $(\Omega^{(1)}, \mathcal{F}^{(1)}, \{\mathcal{F}_{t(i-1)K_n}^{(1)}\}_{i \in \mathbb{N}}, \mathbb{P}^{(1)})$, $H_t^{(1)}$ and $H_t^{(2)}$ are two discrete martingales, $\left\{ \frac{1}{\sqrt{r_n - s_n + 1}} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \right\}_{i \in \mathbb{N}^+}$, and $\left\{ \frac{1}{\sqrt{r_n - s_n + 1}} (m_i m_{i+s_n-1}) \right\}_{i \in \mathbb{N}^+}$ are two martingale difference sequences.

More specifically, the summands in $H_t^{(1)}$ and $H_t^{(2)}$:

$$\begin{aligned}
& \left\{ \frac{1}{\sqrt{r_n - s_n + 1}} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \right\}_{n \in \mathbb{N}^+, i \in \mathbb{N}^+ \leq r_n - s_n + 1} \\
& \left\{ \frac{1}{\sqrt{r_n - s_n + 1}} (m_i m_{i+s_n-1}) \right\}_{n \in \mathbb{N}^+, i \in \mathbb{N}^+ \leq r_n - s_n + 1}
\end{aligned}$$

are two triangular sequences to which we can apply martingale central limit theorem.

Note that by the results of **Lemma 4**, we can get

$$\begin{aligned}
\langle H^{(1)}, H^{(1)} \rangle_T | \omega^{(0)} &= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} E \left[(m_i^2 - E(m_i^2 | \omega^{(0)}))^2 | \mathcal{F}_{t(i-1)K}^{(1)} \right] | \omega^{(0)} \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left[E(m_i^4 | \mathcal{F}_{t(i-1)K}^{(1)}) - \left(E(m_i^2 | \omega^{(0)}, | \mathcal{F}_{t(i-1)K}^{(1)}) \right)^2 \right] | \omega^{(0)} \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left[E(m_i^4 | \omega^{(0)}) - \left(E(m_i^2 | \omega^{(0)}) \right)^2 \right] \\
&= 6 \left[\left(E(\epsilon^4 | \omega^{(0)}) \right)^2 - E(\epsilon^4 | \omega^{(0)}) \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 + \left(E(\epsilon^2 | \omega^{(0)}) \right)^4 \right] - \left(E(\epsilon^4 | \omega^{(0)}) \right)^2 \\
&= 5 \left(E(\epsilon^4 | \omega^{(0)}) \right)^2 - 6 E(\epsilon^4 | \omega^{(0)}) \left(E(\epsilon^2 | \omega^{(0)}) \right)^2 + 6 \left(E(\epsilon^2 | \omega^{(0)}) \right)^4
\end{aligned}$$

And

$$\begin{aligned}
\langle H^{(2)}, H^{(2)} \rangle_T | \omega^{(0)} &= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} E \left[m_i^2 m_{i+s_n-1}^2 | \mathcal{F}_{t(i+s_n-2)K}^{(1)} \right] | \omega^{(0)} \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) + E(m_i^2 | \omega^{(0)}) \right) E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} E(m_i^2 | \omega^{(0)}) \cdot E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) \\
&\quad + \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \cdot E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right)
\end{aligned}$$

where $\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \cdot E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) = o_p(1)$ since $P(\tau_l > T) \xrightarrow{\mathbb{P}^{(0)}} 1$ as $l \rightarrow \infty$ and

$$\begin{aligned}
&E \left(\left(\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \cdot E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) \mathbf{1}_{\{\tau_l > T\}} \right)^2 | \omega^{(0)} \right) \\
&= \frac{1}{(r_n - s_n + 1)^2} \sum_{k=1}^{r_n - s_n + 1} \text{Var} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \left(E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) \right)^2 \mathbf{1}_{\{\tau_l > T\}} \\
&\leq \frac{1}{(r_n - s_n + 1)^2} \sum_{k=1}^{r_n - s_n + 1} M_{(4,l)} \cdot M_{(2,l)}^2 = O_p \left(\frac{1}{r_n - s_n + 1} \right)
\end{aligned}$$

by **Proposition 1**, we know

$$\frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} \left(m_i^2 - E(m_i^2 | \omega^{(0)}) \right) \cdot E \left(m_{i+s_n-1}^2 | \omega^{(0)} \right) \xrightarrow{\mathbb{P}} 0$$

Thus, we have

$$\langle H^{(2)}, H^{(2)} \rangle_T | \omega^{(0)} \rightarrow \left(E(\epsilon^4 | \omega^{(0)}) \right)^2$$

Besides,

$$\begin{aligned}
\langle H^{(1)}, H^{(2)} \rangle_T | \omega^{(0)} &= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} E \left[(m_i^2 - E(m_i^2 | \omega^{(0)})) \cdot m_i m_{i+s_n-1} | \mathcal{F}_{t(i+s_n-2)K}^{(1)} \right] | \omega^{(0)} \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} (m_i^2 - E(m_i^2 | \omega^{(0)})) \cdot m_i E \left(m_{i+s_n-1} | \mathcal{F}_{t(i+s_n-2)K}^{(1)} \right) | \omega^{(0)} \\
&= \frac{1}{r_n - s_n + 1} \sum_{i=1}^{r_n - s_n + 1} (m_i^2 - E(m_i^2 | \omega^{(0)})) m_i \cdot \underbrace{E \left(m_{i+s_n-1} | \omega^{(0)} \right)}_0 = 0
\end{aligned}$$

Therefore, we have the following joint asymptotic distribution for $H_T^{(1)}$ and $H_T^{(2)}$:

$$\begin{pmatrix} H_T^{(1)} \\ H_T^{(2)} \end{pmatrix} \xrightarrow{\mathcal{L}} \mathcal{MN} \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \zeta^2 & 0 \\ 0 & (E(\epsilon^4 | \omega^{(0)}))^2 \end{pmatrix} \right) \quad (71)$$

where $\zeta = \sqrt{5(E(\epsilon^4 | \omega^{(0)}))^2 - 6E(\epsilon^4 | \omega^{(0)}) (E(\epsilon^2 | \omega^{(0)}))^2 + 6(E(\epsilon^2 | \omega^{(0)}))^4}$.

Lastly, note that $R_T = o_p(1)$, because $P(\tau_l > T) \xrightarrow{\mathbb{P}^{(0)}} 1$ as $l \rightarrow \infty$, and

$$\begin{aligned}
E(R_T^2 \mathbf{1}_{\{\tau_l > T\}} | \omega^{(0)}) &= \frac{1}{r_n - s_n + 1} E \left[\left(\sum_{i=r_n-s_n+2}^{r_n} (m_i^2 - E(\epsilon^4 | \omega^{(0)})) - \sum_{i=1}^{s_n-1} (m_i^2 - E(\epsilon^4 | \omega^{(0)})) \right)^2 | \omega^{(0)} \right] \\
&= \frac{1}{r_n - s_n + 1} \left[\sum_{i=r_n-s_n+2}^{r_n} \left(E(m_i^4 | \omega^{(0)}) - (E(\epsilon^4 | \omega^{(0)}))^2 \right) + \sum_{i=1}^{s_n-1} \left(E(m_i^4 | \omega^{(0)}) - (E(\epsilon^4 | \omega^{(0)}))^2 \right) \right] \\
&= O_p \left(\frac{s_n}{r_n - s_n + 1} \right) = o_p(1)
\end{aligned}$$

Plug in this result into (70), we can get:

$$\begin{aligned} \sqrt{r_n - s_n + 1} \left(V(Y, K_n, 2)^n - 2E(\epsilon^4|X) \right) &= 2 \left(H_T^{(1)} + H_T^{(2)} \right) + o_p(1) \\ &\xrightarrow{\mathcal{L}-s} \mathcal{MN} \left(0, \eta^2 \right) \end{aligned} \quad (72)$$

where $\eta = 2\sqrt{6} \cdot \sqrt{(E(\epsilon^4|\omega^{(0)}))^2 - E(\epsilon^4|\omega^{(0)}) (E(\epsilon^2|\omega^{(0)}))^2 + (E(\epsilon^2|\omega^{(0)}))^4}$.

□

11.8 Proof of Theorem 4

The proof for the Theorem 4 is almost the same argument as the proof of Theorem 3 in section 11.7 (the difference is that we do not need to deal with the error term R_T).

11.9 Proof of Theorem 5

Proof. In this proof, we write K and r without the subscript n in order to avoid clustered notation, however, K and r still depend on n .

11.9.1 The law of large number: the limit quantity

Under the assumption of Theorem 5, and follow from Lemma 1, we know:

$$\begin{aligned} \frac{1}{2K} [Y, Y]_{(T_{i-1}, T_i]}^{(all)} &= \frac{1}{2K} [\epsilon, \epsilon]_{(T_{i-1}, T_i]}^{(all)} + O_p \left(\frac{1}{K} \right) \\ &= \frac{1}{K} \sum_{t_j \in (T_{i-1}, T_i]} g_{t_j}(\omega^{(0)}) + \frac{1}{\sqrt{K}} \left(M_{(T_{i-1}, T_i]}^{(1)} - M_{(T_{i-1}, T_i]}^{(2)} \right) + O_p \left(\frac{1}{K} \right) \\ &= \frac{1}{K} \sum_{t_j \in (T_{i-1}, T_i]} g_{t_j}(\omega^{(0)}) + O_p \left(\frac{1}{\sqrt{K}} \right) \end{aligned}$$

where $M_{(T_{i-1}, T_i]}^{(1)}$ and $M_{(T_{i-1}, T_i]}^{(2)}$ are defined analogically as in (8). So we have:

$$\begin{aligned} \frac{1}{2K} [Y, Y]_{(T_i, T_{i+1}]}^{(all)} - \frac{1}{2K} [Y, Y]_{(T_{i-1}, T_i]}^{(all)} &= \frac{1}{K} \sum_{t_j \in (T_{i-1}, T_i]} \left(g_{t_j+K}(\omega^{(0)}) - g_{t_j}(\omega^{(0)}) \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \\ &= \frac{1}{K} \sum_{t_j \in (T_{i-1}, T_i]} \sum_{l=1}^K \left(g_{t_j+l}(\omega^{(0)}) - g_{t_j+l-1}(\omega^{(0)}) \right) + O_p \left(\frac{1}{\sqrt{K}} \right) \\ &= \underbrace{\frac{1}{K} \sum_{j=1}^K \sum_{l=1}^K \left(g_{t_{(i-1)K+j+l}}(\omega^{(0)}) - g_{t_{(i-1)K+j+l-1}}(\omega^{(0)}) \right)}_{(A)} + O_p \left(\frac{1}{\sqrt{K}} \right) \end{aligned}$$

notice that:

$$\begin{aligned} (A)^2 &= \frac{1}{K^2} \left[\sum_{j=1}^K (j-1) \Delta g_{t_{(i-1)K+j}}(\omega^{(0)}) + \sum_{j=K+1}^{2K} (K - (j-1)) \Delta g_{t_{(i-1)K+j}}(\omega^{(0)}) \right]^2 \\ &= \sum_{j=1}^K \frac{(j-1)^2}{K^2} (\Delta g_{t_{(i-1)K+j}}(\omega^{(0)}))^2 + \sum_{j=K+1}^{2K} \frac{(K - (j-1))^2}{K^2} (\Delta g_{t_{(i-1)K+j}}(\omega^{(0)}))^2 + (I) + (II) + (III) \end{aligned}$$

where

$$(I) = \sum_{j=1}^K \sum_{l \neq j} \frac{(j-1)(l-1)}{K^2} \Delta g_{t_{(i-1)K+j}}(\omega^{(0)}) \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \quad (73)$$

$$(II) = \sum_{j=1}^K \sum_{l \neq j} \frac{(K-(j-1))(K-(l-1))}{K^2} \Delta g_{t_{iK+j}}(\omega^{(0)}) \Delta g_{t_{iK+l}}(\omega^{(0)}) \quad (74)$$

$$(III) = \sum_{j=1}^K \sum_{l=1}^K \frac{(j-1)(K-(l-1))}{K^2} \Delta g_{t_{(i-1)K+j}}(\omega^{(0)}) \Delta g_{t_{iK+l}}(\omega^{(0)}) \quad (75)$$

are mean-0 martingales. Since the volatility process for $g_t(\omega^{(0)})$ is locally bounded, by standard localization procedure, we can strengthen the condition by assuming $\sigma_t^{(g)} \leq \sigma_+^{(g)}, \forall t \in [0, T]$, therefore,

$$E[(I)^2] \leq \frac{T^2(\sigma_+^{(g)})^4}{n^2} \sum_{j=1}^K \sum_{l=1}^K \left[\frac{(j-1)(l-1)}{K^2} \right]^2 = \frac{T^2(\sigma_+^{(g)})^4}{n^2} \cdot \sum_{j=1}^K \frac{(j-1)^2}{K^2} \cdot \sum_{l=1}^K \frac{(l-1)^2}{K^2} = O_p \left(\frac{K^2}{n^2} \right)$$

by Chebyshev inequality, $(I) = O_p \left(\frac{K}{n} \right)$. Similarly, $(II), (III) = O_p \left(\frac{K}{n} \right)$. Furthermore, we can know $(A) = O_p \left(\sqrt{\frac{K}{n}} \right)$.

$$\begin{aligned} \text{Thus, } & \sum_{i=1}^{r-1} \left(\frac{1}{2K} [Y, Y]_{(T_i, T_{i+1})}^{(all)} - \frac{1}{2K} [Y, Y]_{(T_{i-1}, T_i)}^{(all)} \right)^2 \\ &= \sum_{j=1}^K \frac{(j-1)^2}{K^2} (\Delta g_{t_j}(\omega^{(0)}))^2 + \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} (\Delta g_{t_{(r-1)K+j}}(\omega^{(0)}))^2 \\ &+ \sum_{i=2}^{r-2} \sum_{j=1}^K \frac{(j-1)^2 + (K-(j-1))^2}{K^2} (\Delta g_{t_{(i-1)K+j}}(\omega^{(0)}))^2 + \underbrace{\sum_{i=1}^{r-1} [(I) + (II) + (III)]}_{O_p \left(\frac{\sqrt{r}K}{n} \right) = O_p \left(\frac{1}{\sqrt{r}} \right)} + \underbrace{O_p \left(\frac{r}{K} \right)}_{\text{error due to noises}} \end{aligned}$$

note that the error due to noises (of the stochastic order $O_p \left(\frac{r}{K} \right)$) approximately equals to $\frac{2r}{TK} \int_0^T h_t(\omega^{(0)}) dt$ according to the **Theorem 3**. Hence,

$$\begin{aligned} & \sum_{i=1}^{r-1} \left(\frac{1}{2K} [Y, Y]_{(T_i, T_{i+1})}^{(all)} - \frac{1}{2K} [Y, Y]_{(T_{i-1}, T_i)}^{(all)} \right)^2 - \frac{2}{3} \sum_{j=1}^n (\Delta g_{t_j}(\omega^{(0)}))^2 - \frac{2r}{TK} \int_0^T h_t(\omega^{(0)}) dt \\ &= \sum_{i=1}^{r-2} \sum_{j=1}^K \left[\frac{1}{3} - \frac{K-(j-1)}{K} \frac{j-1}{K} \right] (\Delta g_{t_{iK+j}}(\omega^{(0)}))^2 \\ &+ \sum_{j=1}^K \left[\frac{(j-1)^2}{K^2} - \frac{2}{3} \right] (\Delta g_{t_j}(\omega^{(0)}))^2 + \sum_{j=1}^K \left[\frac{(K-(j-1))^2}{K^2} - \frac{2}{3} \right] (\Delta g_{t_{(r-1)K+j}}(\omega^{(0)}))^2 + \underbrace{(E)}_{\text{error terms}} \end{aligned}$$

where $(E) = (E1) + (E2)$. $(E1) = O_p \left(\frac{1}{\sqrt{r}} \right)$ is the error due to discretization, $(E2) = O_p \left(\frac{r}{K} \right)$ is the error coming from noises. Moreover,

$$\begin{aligned} & \sum_{i=1}^{r-2} \sum_{j=1}^K \left[\frac{1}{3} - \frac{K-(j-1)}{K} \frac{j-1}{K} \right] (\Delta g_{t_{iK+j}}(\omega^{(0)}))^2 + \sum_{j=1}^K \left[\frac{(j-1)^2}{K^2} - \frac{2}{3} \right] (\Delta g_{t_j}(\omega^{(0)}))^2 \\ &+ \sum_{j=1}^K \left[\frac{(K-(j-1))^2}{K^2} - \frac{2}{3} \right] (\Delta g_{t_{(r-1)K+j}}(\omega^{(0)}))^2 = O_p \left(\frac{K}{n} \right) = O_p \left(\frac{1}{r} \right) \end{aligned}$$

so

$$\sum_{i=1}^{r-1} \left(\frac{1}{2K} [Y, Y]_{(T_i, T_{i+1})}^{(all)} - \frac{1}{2K} [Y, Y]_{(T_{i-1}, T_i)}^{(all)} \right)^2 - \frac{2}{3} \sum_{j=1}^n (\Delta g_{t_j}(\omega^{(0)}))^2 - \frac{2r}{TK} \int_0^T h_t(\omega^{(0)}) dt = (E1) + (E2) + O_p \left(\frac{1}{r} \right)$$

11.9.2 Decomposition of the error process

Because $\frac{2}{3} \sum_{j=1}^n (\Delta g_{t_j}(\omega^{(0)}))^2 - \frac{2}{3} \langle g, g \rangle_T = O_p\left(\frac{1}{\sqrt{n}}\right)^{13}$, we can see

$$\frac{r}{K} U(Y, K, 2)_T^2 - \frac{2}{3} \langle g, g \rangle_T = (\text{E1}) + (\text{E2}) + O_p\left(\frac{1}{r}\right) \quad (76)$$

Since the error (E2) comes from the negligible remaining of the microstructure noise, from the proof of **Theorem 3**, we know $\frac{K}{r} \times (\text{E2})$ would converge to $\frac{2}{T} \int_0^T h_t(\omega^{(0)}) dt$ in probability as $n \rightarrow \infty$ with rate \sqrt{r} :

$$\frac{K}{r} \cdot \sqrt{r} \left((\text{E2}) - \frac{2r}{TK} \int_0^T h_t(\omega^{(0)}) dt \right) \xrightarrow{\mathcal{L}-s} \mathcal{MN}\left(0, \tilde{\eta}^2\right)$$

where $\tilde{\eta}^2 = \frac{24}{T} \int_0^T \left[h_t^2(\omega^{(0)}) - h_t(\omega^{(0)}) \cdot g_t^2(\omega^{(0)}) + g_t^4(\omega^{(0)}) \right] dt$.

On the other hand, following the previous calculation (especially (73)),

$$\begin{aligned} (\text{E1}) &= 2 \sum_{i=1}^{r-1} \sum_{j=2}^K \sum_{l=1}^{j-1} \frac{1}{K^2} [(j-1)(l-1) + (K-(j-1))(K-(l-1))] \cdot \Delta g_{t_{iK+j}}(\omega^{(0)}) \Delta g_{t_{iK+l}}(\omega^{(0)}) \\ &\quad + \sum_{i=1}^{r-1} \sum_{j=1}^K \sum_{l=1}^K \frac{1}{K^2} (l-1)(K-(j-1)) \cdot \Delta g_{t_{iK+j}}(\omega^{(0)}) \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) + \underbrace{O_p\left(\frac{K}{n}\right)}_{\text{the edge in (E1)}} \end{aligned}$$

if we define the following two quantities:

$$N_T^{(1)} \equiv 2\sqrt{r} \sum_{i=0}^{r-1} \sum_{j=2}^K \Delta g_{t_{iK+j}}(\omega^{(0)}) \cdot \left[\sum_{l=1}^{j-1} \left(1 + 2 \frac{(j-1)(l-1)}{K^2} - \frac{j-1}{K} - \frac{l-1}{K} \right) \Delta g_{t_{iK+l}}(\omega^{(0)}) \right] \quad (77)$$

$$N_T^{(2)} \equiv \sqrt{r} \sum_{i=1}^{r-1} \sum_{j=1}^K \left(1 - \frac{j-1}{K} \right) \Delta g_{t_{iK+j}}(\omega^{(0)}) \cdot \left(\sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \right) \quad (78)$$

then we have:

$$(\text{E1}) = \frac{1}{\sqrt{r}} N_T^{(1)} + \frac{1}{\sqrt{r}} N_T^{(2)} + O_p\left(\frac{1}{r}\right) \quad (79)$$

Furthermore, by (77)

$$\begin{aligned} \langle N^{(1)}, N^{(2)} \rangle_T &= 2r \sum_{i=1}^{r-1} \sum_{j=2}^K \left(1 - \frac{j-1}{K} \right) \Delta \langle g, g \rangle_{t_{iK+j}} \times \sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \\ &\quad \times \sum_{l=1}^{j-1} \left(1 + 2 \frac{(j-1)(l-1)}{K^2} - \frac{j-1}{K} - \frac{l-1}{K} \right) \Delta g_{t_{iK+l}}(\omega^{(0)}) \end{aligned}$$

$$\begin{aligned} E \left[\langle N^{(1)}, N^{(2)} \rangle_T^2 \right] &= 4r^2 \sum_{i=1}^{r-1} \sum_{j=2}^K \left(1 - \frac{j-1}{K} \right)^2 (\Delta \langle g, g \rangle_{t_{iK+j}})^2 \times \left(\sum_{l=1}^K \frac{(l-1)^2}{K^2} E \left[\left(\Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \right)^2 \right] \right) \\ &\quad \times \left(\sum_{l=1}^{j-1} \left(\frac{(j-1)(l-1)}{K^2} + \frac{(K-(j-1))(K-(l-1))}{K^2} \right)^2 E \left[\left(\Delta g_{t_{iK+l}}(\omega^{(0)}) \right)^2 \right] \right) \end{aligned}$$

note that $\sum_{l=1}^K \frac{(l-1)^2}{K^2} E \left[\left(\Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \right)^2 \right] = O_p\left(\frac{K}{n}\right)$, and

¹³The proofs can be found in Jacod and Protter [1998], Mykland and Zhang [2006].

$$\begin{aligned}
& \sum_{l=1}^{j-1} \left(\frac{(j-1)(l-1)}{K^2} + \frac{(K-(j-1))(K-(l-1))}{K^2} \right)^2 E \left[\left(\Delta g_{t_{iK+l}}(\omega^{(0)}) \right)^2 \right] \\
&= \sum_{l=1}^{j-1} \left[\frac{(j-1) - (K-(j-1))}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 E \left[\left(\Delta g_{t_{iK+l}}(\omega^{(0)}) \right)^2 \right] \\
&= \sum_{l=1}^{j-1} \left[\frac{(2(j-1)-K)^2}{K^4} (l-1)^2 + \frac{2(K-(j-1))(2(j-1)-K)}{K^3} (l-1) + \frac{(K-(j-1))^2}{K^2} \right] E \left[\left(\Delta g_{t_{iK+l}}(\omega^{(0)}) \right)^2 \right] \\
&= O \left(\frac{K^2 j^3 + K j^4 + j^5}{K^4} + \frac{K^2 j^2 + K j^3 + j^4}{K^3} + \frac{K^2 j + K j^2 + j^3}{K^2} \right) \times O_p \left(\frac{1}{n} \right)
\end{aligned}$$

thus we can know $E \left[\langle N^{(1)}, N^{(2)} \rangle_T^2 \right] = O_p \left(\frac{r^3 K^3}{n^4} \right) = O_p \left(\frac{1}{\sqrt{n}} \right) \xrightarrow{\mathbb{P}^{(0)}} 0$. So

$$\langle N^{(1)} + N^{(2)}, N^{(1)} + N^{(2)} \rangle_T = \langle N^{(1)}, N^{(1)} \rangle_T + \langle N^{(2)}, N^{(2)} \rangle_T + O_p \left(\frac{1}{\sqrt{n}} \right) \quad (80)$$

11.9.3 Calculating $\langle N^{(1)}, N^{(1)} \rangle_T$

By (77),

$$\begin{aligned}
\langle N^{(1)}, N^{(1)} \rangle_T &= 4r \sum_{i=0}^{r-1} \sum_{j=2}^K \Delta \langle g, g \rangle_{t_{iK+j}} \times \left[\sum_{l=1}^{j-1} \left(\frac{(j-1)(l-1)}{K^2} + \frac{(K-(j-1))(K-(l-1))}{K^2} \right) \Delta g_{t_{iK+l}} \right]^2 \\
&= (A1) + (A2)
\end{aligned}$$

where

$$\begin{aligned}
(A1) &= 4r \sum_{i=0}^{r-1} \sum_{j=2}^K \Delta \langle g, g \rangle_{t_{iK+j}} \times \left[\sum_{l=1}^{j-1} \left(\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right)^2 (\Delta g_{t_{iK+l}})^2 \right] \\
&= 4r \sum_{i=0}^{r-1} \sum_{j=2}^K \left(\sigma_{t_{iK+j}}^{(g)} \right)^4 \Delta_n^2 \times \sum_{l=1}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 + O_p \left(\frac{1}{rn} \right)
\end{aligned}$$

the error term appears because $\sigma^{(g)}$ is an Itô process, and error due to the local-consistency approximation for $\sigma^{(g)}$ is of the same order as

$$4r \sum_{i=0}^{r-1} \sum_{j=2}^K j \Delta_n^4 \sum_{l=1}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 = O_p \left(\frac{1}{rn} \right)$$

Besides,

$$(A2) = 8r \sum_{i=0}^{r-1} \sum_{j=3}^K \Delta \langle g, g \rangle_{t_{iK+j}} \cdot \phi_j$$

where

$$\phi_j = \sum_{l=2}^{j-1} \sum_{k=1}^{l-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right] \cdot \left[\frac{2(j-1)-K}{K^2} (k-1) + \frac{K-(j-1)}{K} \right] \Delta g_{t_{iK+l}}(\omega^{(0)}) \Delta g_{t_{iK+k}}(\omega^{(0)})$$

by Burkholder-Davis-Gundy inequality on ϕ_j , $\exists C_1 \in \mathbb{R}^+$ such that $\|\phi_j\|_2^2 \leq C_1^2 \|\langle \phi_j, \phi_j \rangle\|_1$

$$\begin{aligned}
&= C_1^2 E \sum_{l=2}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 \Delta \langle g, g \rangle_{t_{iK+l}} \cdot \left[\sum_{k=1}^{l-1} \left(\frac{2(j-1)-K}{K^2} (k-1) + \frac{K-(j-1)}{K} \right) \Delta g_{t_{iK+k}}(\omega^{(0)}) \right]^2 \\
&\leq \frac{C_1^2 \left(\sigma_+^{(g)} \right)^2}{n^2} \sum_{l=2}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 \times \sum_{k=1}^{l-1} \left[\frac{2(j-1)-K}{K^2} (k-1) + \frac{K-(j-1)}{K} \right]^2
\end{aligned}$$

$$\begin{aligned}
& \text{notice that } \sum_{k=1}^{l-1} \left[\frac{2(j-1)-K}{K^2} (k-1) + \frac{K-(j-1)}{K} \right]^2 \\
&= \sum_{k=1}^{l-1} \left[\frac{(2(j-1)-K)^2}{K^4} (k-1)^2 + \frac{2(K-(j-1))(2(j-1)-K)}{K^3} (k-1) + \frac{(K-(j-1))^2}{K^2} \right] \\
&= O_p \left(\frac{K^2 + Kj + J^2}{K^4} l^3 + \frac{K^2 + Kj + j^2}{K^3} l^2 + \frac{K^2 + Kj + j^2}{K^2} l \right)
\end{aligned}$$

$$\begin{aligned}
& \text{so } \sum_{l=2}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 \times \sum_{k=1}^{l-1} \left[\frac{2(j-1)-K}{K^2} (k-1) + \frac{K-(j-1)}{K} \right]^2 \\
&= O \left(\sum_{l=2}^{j-1} \left(\frac{K^2 + Kj + j^2}{K^4} l^2 + \frac{K^2 + Kj + j^2}{K^3} l + \frac{K^2 + Kj + j^2}{K^2} \right)^2 \times l \right)
\end{aligned}$$

$$\text{thus, } \|\phi_j\|_2 = O_p \left(\frac{j^3}{K^2 n} \right) \text{ and } \sum_{j=3}^K \|\phi_j\|_2^2 \leq \sum_{j=3}^K C_1^2 \|\langle \phi_j, \phi_j \rangle\|_1 = O_p \left(\frac{K^3}{n^2} \right).$$

Define $(A2)' \equiv 8r \sum_{i=0}^{r-1} \sum_{j=3}^K \left(\sigma_{t_{iK}}^{(g)} \right)^2 \Delta_n \phi_j$, and apply Burkholder-Davis-Gundy inequality again, but on (A2), we get:

$$E[(A2)]^2 = \|(A2)\|_2^2 \leq 64r^2 C_2^2 \sum_{i=0}^{r-1} \sum_{j=3}^K \left(\sigma_{t_{iK}}^{(g)} \right)^4 \Delta_n^2 \times \|\langle \phi_j, \phi_j \rangle\|_1 = O_p \left(\frac{r^2 K^3}{n^4} \right) = O_p \left(\frac{1}{n} \right)$$

so

$$(A2)' = O_p \left(\frac{1}{\sqrt{n}} \right) \quad (81)$$

Then by Cauchy-Schwarz inequality,

$$\begin{aligned}
\|(A2) - (A2)'\|_1 &= \left\| 8r \sum_{i=0}^{r-1} \sum_{j=3}^K \left[\Delta \langle g, g \rangle_{t_{iK+j}} - \left(\sigma_{t_{iK}}^{(g)} \right)^2 \Delta_n \right] \phi_j \right\|_1 \leq 8r \sum_{i=0}^{r-1} \sum_{j=3}^K \left\| \left[\Delta \langle g, g \rangle_{t_{iK+j}} - \left(\sigma_{t_{iK}}^{(g)} \right)^2 \Delta_n \right] \phi_j \right\|_1 \\
&\leq 8r \sum_{i=0}^{r-1} \sum_{j=3}^K \left\| \Delta \langle g, g \rangle_{t_{iK+j}} - \left(\sigma_{t_{iK}}^{(g)} \right)^2 \Delta_n \right\| \cdot \|\phi_j\|_2 \\
&\leq 8r^2 \Delta(\mathcal{G}) \sum_{j=3}^K \left\| \sup_{|t-s| \leq K \Delta(\mathcal{G})} \left[\left(\sigma_t^{(g)} \right)^2 - \left(\sigma_s^{(g)} \right)^2 \right] \right\|_2 \cdot \|\phi_j\|_2 \\
&\leq 8r^2 \sqrt{K} (\Delta(\mathcal{G}))^{\frac{3}{2}} \sum_{j=3}^K \|\phi_j\|_2 \\
&= O_p \left(\frac{r^2 K^{\frac{5}{2}}}{n^{\frac{5}{2}}} \right)
\end{aligned}$$

so

$$\|(A2) - (A2)'\|_1 = O_p \left(\frac{1}{\sqrt{r}} \right) \quad (82)$$

from (81) and (82), we can know $(A2) = o_p(1)$, and more importantly,

$$\langle N^{(1)}, N^{(1)} \rangle_T = 4r \sum_{i=0}^{r-1} \sum_{j=2}^K \left(\sigma_{t_{iK+j}}^{(g)} \right)^4 \Delta_n^2 \times \underbrace{\sum_{l=1}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2}_{(1)} + o_p(1)$$

notice that

$$\begin{aligned}
(1) &= \sum_{l=1}^{j-1} \left[\frac{(2(j-1)-K)^2}{K^4} (l-1)^2 + 2 \frac{(2(j-1)-K)(K-(j-1))}{K^3} (l-1) + \frac{(K-(j-1))^2}{K^2} \right] \\
&= \frac{(2(j-1)-K)^2}{3K^4} j^3 + \frac{(2(j-1)-K)(K-(j-1))}{K^3} j^2 + \frac{(K-(j-1))^2}{K^2} j + O(1) \\
&= \frac{4}{3} \frac{j^5}{K^4} - \frac{10}{3} \frac{j^4}{K^3} + \frac{13}{3} \frac{j^3}{K^2} - 3 \frac{j^2}{K} + j + O(1)
\end{aligned}$$

so

$$\begin{aligned}
\langle N^{(1)}, N^{(1)} \rangle_T &= 4r \sum_{i=0}^{r-1} \sum_{j=2}^K (1) \cdot \left(\sigma^{(g)}_{t_{iK+j}} \right)^4 \Delta_n^2 + o_p(1) \\
&= 4r \sum_{i=0}^{r-1} \sum_{j=2}^K (1) \cdot \left[\left(\sigma^{(g)}_{t_{iK+j}} \right)^2 + O_p \left(\sqrt{K \Delta_n} \right) \right]^2 \Delta_n^2 + o_p(1) \\
&= 4r \sum_{i=0}^{r-1} \left[\left(\sum_{j=2}^K (1) \right) \times \frac{\Delta_n}{K} \left(\sigma^{(g)}_{t_{iK+j}} \right)^4 K \Delta_n + \left(\sum_{j=2}^K (1) \right) \times O_p \left(K^{\frac{1}{2}} \Delta_n^{\frac{5}{2}} \right) \right] + o_p(1) \\
&= 4r \sum_{i=0}^{r-1} \left[\left(\sum_{j=2}^K (1) \right) \times \frac{\Delta_n}{K} \left(\sigma^{(g)}_{t_{iK+j}} \right)^4 K \Delta_n \right] + o_p(1)
\end{aligned}$$

By Faulhaber's formula, we know

$$\sum_{j=2}^K (1) = \left(\frac{4}{3} \times \frac{1}{6} - \frac{10}{3} \times \frac{1}{5} + \frac{13}{3} \times \frac{1}{4} - 3 \times \frac{1}{3} + \frac{1}{2} \right) K^2 + O(K) = \frac{5}{36} K^2 + O(K)$$

so

$$\langle N^{(1)}, N^{(1)} \rangle_T = T \sum_{i=1}^{r-1} \left[\frac{5}{9} + O_p \left(\frac{1}{K} \right) \right] \left(\sigma^{(g)}_{t_{iK}} \right)^4 K \Delta_n + o_p(1) \longrightarrow \frac{5T}{9} \int_0^T \left(\sigma_t^{(g)} \right)^4 dt$$

11.9.4 Calculating $\langle N^{(2)}, N^{(2)} \rangle_T$

By (77),

$$\begin{aligned}
\langle N^{(2)}, N^{(2)} \rangle_T &= r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K - (j-1))^2}{K^2} \Delta \langle g, g \rangle_{t_{iK+j}} \times \left(\sum_{l=1}^K \frac{(l-1)}{K} \Delta g_{t_{(i-1)K+l}} (\omega^{(0)}) \right)^2 \\
&= (B1) + (B2)
\end{aligned}$$

where

$$\begin{aligned}
(B1) &= r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K - (j-1))^2}{K^2} \Delta \langle g, g \rangle_{t_{iK+j}} \times \sum_{l=1}^K \frac{(l-1)^2}{K^2} (\Delta g_{t_{(i-1)K+l}} (\omega^{(0)}))^2 \\
&= r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K - (j-1))^2}{K^2} \left(\sigma^{(g)}_{t_{iK+j}} \right)^4 \Delta_n^2 \times \sum_{l=1}^K \frac{(l-1)^2}{K^2} + O_p \left(\frac{1}{r} \right)
\end{aligned}$$

the error term just above comes from the local-constancy approximation on $(\sigma^{(g)})^2$, it is of the stochastic order $O_p \left(\frac{1}{r} \right)$ because

$$r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K - (j-1))^2}{K^2} 2K \Delta_n^3 \times \sum_{l=1}^K \frac{(l-1)^2}{K^2} = O_p \left(\frac{1}{r} \right)$$

Besides,

$$(B2) = 2r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K - (j-1))^2}{K^2} \Delta \langle g, g \rangle_{t_{iK+j}} \psi_i$$

where

$$\psi_i = \sum_{l=2}^K \sum_{k=1}^{l-1} \frac{l-1}{K} \frac{k-1}{K} \Delta g_{t_{(i-1)K+k}} (\omega^{(0)}) \Delta g_{t_{(i-1)K+l}} (\omega^{(0)})$$

Apply Burkholder-Davis-Gundy on ψ_i , since $(\psi_i)_t \equiv \sum_{l=2}^K \sum_{k=1}^{l-1} \frac{l-1}{K} \frac{k-1}{K} \Delta g_{t_{(i-1)K+k}} (\omega^{(0)}) \int_{t_{(i-1)K+l-1}}^{t_{(i-1)K+l}} d g_t (\omega^{(0)})$ is a continuous martingale by assumption of the **Theorem 5**,

$$\begin{aligned}
\|\psi_i\|_2^2 &\leq D_1^2 \|\langle \psi_i, \psi_i \rangle\|_1 = D_1^2 E \sum_{l=2}^K \frac{(l-1)^2}{K^2} \Delta \langle g, g \rangle_{t_{(i-1)K+l}} \times \left(\sum_{k=1}^{l-1} \frac{k-1}{K} \Delta g_{t_{(i-1)K+k}} \right)^2 \\
&\leq D_1^2 \left(\sigma_+^{(g)} \right)^4 (\Delta(g))^2 \times \sum_{l=2}^K \frac{(l-1)^2}{K^2} \sum_{k=1}^{l-1} \frac{(k-1)^2}{K^2} = O_p \left(\left(\frac{K}{n} \right)^2 \right)
\end{aligned}$$

so $\|\psi_i\|_2^2 \leq D_1 \|\langle \psi_i, \psi_i \rangle\|_1 = O_p\left(\frac{1}{r^2}\right)$. Furthermore, define $(B2)' \equiv 2r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \left(\sigma_{t_{(i-1)K}}^{(g)}\right)^2 \Delta_n \psi_i$, apply Burkholder-Davis-Gundy inequality again on $(B2)'$,

$$\begin{aligned} \|(B2)'\|_2^2 &= E[(B2)']^2 \leq 4r^2 D_2^2 \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^4}{K^4} \left(\sigma_t^{(g)}\right)^4 \Delta_n^2 \times \|\langle \psi_i, \psi_i \rangle\|_1 \\ &\leq 4r^2 D_2^2 \left(\sigma_+^{(g)}\right)^4 (\Delta(\mathcal{G}))^2 \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^4}{K^4} \|\langle \psi_i, \psi_i \rangle\|_1 \\ &= O_p\left(\frac{r^2}{n^2}\right) \times \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^4}{K^4} \times O_p\left(\frac{K^2}{n^2}\right) = O_p\left(\frac{1}{n}\right) \end{aligned}$$

therefore,

$$(B2)' = O_p\left(\frac{1}{\sqrt{n}}\right) \quad (83)$$

By Cauchy-Schwarz inequality,

$$\begin{aligned} \|(B2) - (B2)'\|_1 &\leq 2r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \left\| \left[\Delta \langle g, g \rangle_{t_{iK+j}} - \left(\sigma_{t_{(i-1)K}}^{(g)}\right)^2 \Delta_n \right] \psi_i \right\|_1 \\ &\leq 2r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \left\| \Delta \langle g, g \rangle_{t_{iK+j}} - \left(\sigma_{t_{(i-1)K}}^{(g)}\right)^2 \Delta_n \right\|_2 \cdot \|\psi_i\|_2 \\ &\leq 2r^2 K \Delta(\mathcal{G}) \cdot \left\| \sup_{|t-s| \leq 2K\Delta(\mathcal{G})} \left[\left(\sigma_t^{(g)}\right)^2 - \left(\sigma_s^{(g)}\right)^2 \right] \right\|_2 \cdot \sup_i \|\psi_i\|_2 \end{aligned}$$

so

$$\|(B2) - (B2)'\|_1 = O_p\left(\frac{1}{\sqrt{r}}\right) \quad (84)$$

combine (83) and (84), we can get $(B2) = o_p(1)$. More importantly,

$$\begin{aligned} \langle N^{(2)}, N^{(2)} \rangle_T &= r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \left(\sigma_{t_{iK+j}}^{(g)}\right)^4 \Delta_n^2 \times \sum_{l=1}^K \frac{(l-1)^2}{K^2} + o_p(1) \\ &= r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \left[\left(\sigma_{t_{iK}}^{(g)}\right)^4 + O_p(\sqrt{K\Delta_n}) \right] \times \left(\frac{K}{3} + O_p(1) \right) \times \Delta_n^2 + o_p(1) \\ &= r \sum_{i=1}^{r-1} \sum_{j=1}^K \left(\frac{K}{3} + O_p(1) \right) \times \left(\sigma_{t_{iK}}^{(g)}\right)^4 \times \left(\frac{K}{3} + O_p(1) \right) \times \Delta_n^2 + O_p\left(\frac{1}{\sqrt{r}}\right) + o_p(1) \\ &= r \sum_{i=1}^{r-1} \frac{K^2}{9} \left(\sigma_{t_{iK}}^{(g)}\right)^4 \Delta_n^2 + o_p(1) = T \sum_{i=1}^{r-1} \frac{1}{9} \left(\sigma_{t_{iK}}^{(g)}\right)^4 \cdot K\Delta_n + o_p(1) \\ &\rightarrow \frac{T}{9} \int_0^T \left(\sigma_t^{(g)}\right)^4 dt \end{aligned}$$

11.9.5 Proof of the stable convergence

Following the results of the Subsections 11.9.2, 11.9.3 and 11.9.4,

$$\langle \sqrt{r}(E1), \sqrt{r}(E1) \rangle = \langle N^{(1)} + N^{(2)}, N^{(1)} + N^{(2)} \rangle_T = \langle N^{(1)}, N^{(1)} \rangle_T + \langle N^{(2)}, N^{(2)} \rangle_T = \frac{2T}{3} \int_0^T \left(\sigma_t^{(g)}\right)^4 dt + o_p(1) \quad (85)$$

And following the similar method as that in the proof of Theorem 3 in Subsection 11.7, we know

$$\langle \sqrt{r}(E2), \sqrt{r}(E2) \rangle_T = \frac{24r^2}{TK^2} \int_0^T \left[h_t^2(\omega^{(0)}) - h_t(\omega^{(0)}) g_t^2(\omega^{(0)}) + g_t^4(\omega^{(0)}) \right] dt + o_p\left(\frac{r^2}{K^2}\right) = O_p\left(\frac{r^2}{K^2}\right) \quad (86)$$

We need a technical condition on the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ to which all the relevant processes are adapted:

Condition on the Filtration: there are Brownian motions $W^{(1)}, W^{(2)}, \dots, W^{(p)}$ that generate the filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

Consider the normalized error process (multiply the error process with \sqrt{r}):

$$\begin{aligned} \sqrt{r}(E) &= \sqrt{r}(E1) + \sqrt{r}(E2) = N_T^{(1)} + N_T^{(2)} + \underbrace{O_p\left(\frac{1}{\sqrt{r}}\right)}_{\text{the edge in } N_T^{(1)} + N_T^{(2)}} + \underbrace{\sqrt{r}(E2)}_{O_p\left(\frac{\sqrt{r}}{K}\right)} \\ &= 2\sqrt{r} \sum_{i=0}^{r-1} \sum_{j=2}^K \Delta g_{t_{iK+j}}(\omega^{(0)}) \times \sum_{l=1}^{j-1} \left(1 + 2\frac{j-1}{K} \frac{l-1}{K} - \frac{j-1}{K} - \frac{l-1}{K}\right) \Delta g_{t_{iK+l}}(\omega^{(0)}) \\ &\quad + \sqrt{r} \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{K-(j-1)}{K} \Delta g_{t_{iK+j}}(\omega^{(0)}) \times \sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) + O_p\left(\frac{1}{\sqrt{r}}\right) \end{aligned}$$

Define

$$\begin{aligned} N_t^n &= 2\sqrt{r} \sum_{i=0}^{r-1} \sum_{j=2}^K \left[\sum_{l=1}^{j-1} \left(1 + 2\frac{j-1}{K} \frac{l-1}{K} - \frac{j-1}{K} - \frac{l-1}{K}\right) \Delta g_{t_{iK+l}}(\omega^{(0)}) \right] \cdot \Delta g_{t_{iK+j} \wedge t}(\omega^{(0)}) \\ &\quad + \sqrt{r} \sum_{i=1}^{r-1} \sum_{j=1}^K \left(\sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \right) \cdot \frac{K-(j-1)}{K} \Delta g_{t_{iK+j} \wedge t}(\omega^{(0)}) \end{aligned}$$

then $\sqrt{r}(E1)_t = N_t^n + o_p(1)$.

Suppose $t_{iK+j-1} = \max\{t_k, k = 0, 1, \dots, n, t_k \leq t\}$, then

$$\begin{aligned} d\langle N^n, W^{(i)} \rangle_t &= 2\sqrt{r} \left[\sum_{l=1}^{j-1} \left(1 + 2\frac{j-1}{K} \frac{l-1}{K} - \frac{j-1}{K} - \frac{l-1}{K}\right) \Delta g_{t_{iK+l}}(\omega^{(0)}) \right] d\langle g, W^{(i)} \rangle_t \\ &\quad + \sqrt{r} \left[\sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \cdot \frac{K-(j-1)}{K} \right] d\langle g, W^{(i)} \rangle_t \end{aligned}$$

For $i = 1, 2, \dots, p$, by Kunita-Watanabe inequality,

$$\begin{aligned} \left| \langle g, W^{(i)} \rangle_{t+h} - \langle g, W^{(i)} \rangle_t \right| &\leq \sqrt{\langle g, g \rangle_{t+h} - \langle g, g \rangle_t} \cdot \sqrt{\langle W^{(i)}, W^{(i)} \rangle_{t+h} - \langle W^{(i)}, W^{(i)} \rangle_t} \\ &\leq \sqrt{(\sigma_+^{(g)})^2 h} \cdot \sqrt{h} = \sigma_+^{(g)} h \end{aligned}$$

so $\Delta \langle g, W^{(i)} \rangle_{t_k} \leq \sigma_+^{(g)} \Delta(\mathcal{G})$.

$$\begin{aligned} \langle N^n, W^{(i)} \rangle_T &= 2\sqrt{r} \sum_{i=0}^{r-1} \sum_{j=2}^K \int_{t_{iK+j-1}}^{t_{iK+j}} \sum_{l=1}^{j-1} \left(1 + 2\frac{j-1}{K} \frac{l-1}{K} - \frac{j-1}{K} - \frac{l-1}{K}\right) \Delta g_{t_{iK+l}}(\omega^{(0)}) d\langle g, W^{(i)} \rangle_t \\ &\quad + \sqrt{r} \sum_{i=1}^{r-1} \sum_{j=1}^K \int_{t_{iK+j-1}}^{t_{iK+j}} \frac{K-(j-1)}{K} \sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) d\langle g, W^{(i)} \rangle_t \\ &= \underbrace{2\sqrt{r} \sum_{i=0}^{r-1} \sum_{j=2}^K \left[\sum_{l=1}^{j-1} \frac{(j-1)(l-1) + (K-(j-1))(K-(l-1))}{K^2} \Delta g_{t_{iK+l}}(\omega^{(0)}) \right] \Delta \langle g, W^{(i)} \rangle_{t_{iK+j}}}_{\text{NW1}} \\ &\quad + \underbrace{\sqrt{r} \sum_{i=1}^{r-1} \sum_{j=1}^K \left[\frac{K-(j-1)}{K} \sum_{l=1}^K \frac{l-1}{K} \Delta g_{t_{(i-1)K+l}}(\omega^{(0)}) \right] \Delta \langle g, W^{(i)} \rangle_{t_{iK+j}}}_{\text{NW2}} \end{aligned}$$

note that

$$\begin{aligned}
E(\text{NW1})^2 &\leq \frac{4\left(\sigma_+^{(g)}\right)^6}{n^3} r \sum_{i=0}^{r-1} \sum_{j=2}^K \sum_{l=1}^{j-1} \left[\frac{2(j-1)-K}{K^2} (l-1) + \frac{K-(j-1)}{K} \right]^2 = O_p\left(\frac{r^2 K^2}{n^3}\right) = O_p\left(\frac{1}{n}\right) \\
E(\text{NW2})^2 &\leq \frac{\left(\sigma_+^{(g)}\right)^6}{n^3} r \sum_{i=1}^{r-1} \sum_{j=1}^K \frac{(K-(j-1))^2}{K^2} \sum_{l=1}^K \frac{(l-1)^2}{K^2} = O_p\left(\frac{r^2 K^2}{n^3}\right) = O_p\left(\frac{1}{n}\right)
\end{aligned}$$

the first equality in the first line follows the calculation of $\langle N^{(1)}, N^{(1)} \rangle_T$ in **Subsection 11.9.3**.

Hence, $\langle N^n, W^{(i)} \rangle_T = O_p\left(\frac{1}{\sqrt{n}}\right)$, combine the result for $\langle N^{(1)}, N^{(1)} \rangle_T$ and $\langle N^{(2)}, N^{(2)} \rangle_T$, the **Theorem 5** follows. \square

11.10 Proof of Lemma 5

Proof. Suppose that $E\epsilon_t^2$ is linearly correlated with σ_t^2 and we have

$$E\epsilon_t^2 = \beta\sigma_t^2 + \alpha + \zeta_t \quad (87)$$

for some real numbers β , α and ζ_t is a mean-zero random variable independent of $E\epsilon_t^2$ and σ_t^2 .

From (87), we can get:

$$\widehat{E\epsilon_t^2} + (E\epsilon_t^2 - \widehat{E\epsilon_t^2}) = \beta_n \hat{\sigma}_t^2 + \alpha_n + (\beta\sigma_t^2 - \beta_n \hat{\sigma}_t^2) + (\alpha - \alpha_n) + \zeta_t$$

so, in (43) we have:

$$\begin{aligned}
\eta_t^{(n)} &= \widehat{E\epsilon_t^2} - \beta_n \hat{\sigma}_t^2 - \alpha_n \\
&= (\beta\sigma_t^2 - \beta_n \hat{\sigma}_t^2) + (\alpha - \alpha_n) - (E\epsilon_t^2 - \widehat{E\epsilon_t^2}) + \zeta_t \\
&= \beta \hat{\sigma}_t^2 + \beta(\sigma_t^2 - \hat{\sigma}_t^2) - \beta_n \hat{\sigma}_t^2 + (\alpha - \alpha_n) + (\widehat{E\epsilon_t^2} - E\epsilon_t^2) + \zeta_t
\end{aligned} \quad (88)$$

Note that $\hat{\sigma}_t^2 \rightarrow \sigma_t^2$ and $\widehat{E\epsilon_t^2} \rightarrow E\epsilon_t^2$, thus plug (88) into (43), we get

$$\widehat{E\epsilon_t^2} = \beta \hat{\sigma}_t^2 + \alpha + \zeta_t + o_p(1)$$

so by doing the linear regression on the pairs $(\hat{\sigma}_t^2, \widehat{E\epsilon_t^2})$'s, the slope β_n converges to the real slope β and $\alpha_n \rightarrow \alpha$ in the in-fill asymptotic setting, provided (87) holds. \square